

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 10:13:23 : Search time 373 Seconds  
(without alignments)  
282.341 Million cell updates/sec

Title: US-09-904-603-1  
Perfect score: 625  
Sequence: 1 MSPDRPFKORRSFADRCKEY.....EKDEGFLYMYASQETPGF 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1:	/cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2:	/cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3:	/cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4:	/cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5:	/cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6:	/cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7:	/cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8:	/cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9:	/cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10:	/cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11:	/cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12:	/cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13:	/cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14:	/cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15:	/cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16:	/cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17:	/cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18:	/cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19:	/cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20:	/cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21:	/cgn2_6/ptodata/1/paa/US097A_COMB.pep:*
22:	/cgn2_6/ptodata/1/paa/US097B_COMB.pep:*
23:	/cgn2_6/ptodata/1/paa/US098_COMB.pep:*
24:	/cgn2_6/ptodata/1/paa/US099A_COMB.pep:*
25:	/cgn2_6/ptodata/1/paa/US099B_COMB.pep:*
26:	/cgn2_6/ptodata/1/paa/US100_COMB.pep:*
27:	/cgn2_6/ptodata/1/paa/US101_COMB.pep:*
28:	/cgn2_6/ptodata/1/paa/US102_COMB.pep:*
29:	/cgn2_6/ptodata/1/paa/US103_COMB.pep:*
30:	/cgn2_6/ptodata/1/paa/US104_COMB.pep:*
31:	/cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	625	100.0	121	24	US-09-904-603-1
2	625	100.0	169	1	PCT-US01-18569-2536
					Sequence 1, Appl 1
					Sequence 2536, Ap

3	625	100.0	169	28	US-10-264-049-2536	Sequence 2536, Ap
4	625	100.0	199	1	PCT-US01-08631-40152	Sequence 40152, Ap
5	621	99.4	181	30	US-10-405-027-4118	Sequence 4118, Ap
6	557	89.1	293	31	US-60-212-664-410	Sequence 410, App
7	545	87.2	107	23	US-09-834-366-14523	Sequence 14523, A
8	545	87.2	107	31	US-60-197-873-14523	Sequence 14523, A
9	521	83.4	123	23	US-09-834-366-15699	Sequence 15699, A
10	521	83.4	123	31	US-60-197-873-15699	Sequence 15699, A
11	521	83.4	125	18	US-09-488-725A-2805	Sequence 2805, Ap
12	521	83.4	125	23	US-09-834-366-20036	Sequence 20036, Ap
13	521	83.4	125	28	US-10-219-051B-7991	Sequence 7991, Ap
14	521	83.4	125	28	US-10-219-051B-13087	Sequence 13087, A
15	521	83.4	125	31	US-60-197-873-20036	Sequence 20036, A
16	521	83.4	142	13	US-08-944-467-7	Sequence 7, Appl 1
17	521	83.4	142	24	US-09-904-603-3	Sequence 3, Appl 1
18	521	83.4	142	28	US-10-205-194-71	Sequence 71, Appl 1
19	521	83.4	142	28	US-10-219-051B-13085	Sequence 13085, A
20	521	83.4	155	18	US-09-488-725A-6377	Sequence 6377, Ap
21	521	83.4	161	23	US-09-834-366-24980	Sequence 24980, A
22	521	83.4	161	31	US-60-197-873-24980	Sequence 24980, A
23	521	83.4	171	1	PCT-US00-05882-1051	Sequence 1051, Ap
24	521	83.4	171	24	US-09-925-301-1051	Sequence 1051, Ap
25	516	82.6	119	22	US-09-791-537-111211	Sequence 111211, A
26	516	82.6	141	28	US-10-219-051B-7989	Sequence 7989, Ap
27	466.5	74.6	99	1	PCT-US01-08631-40153	Sequence 40153, A
28	433	69.3	88	23	US-09-834-366-25917	Sequence 25917, A
29	433	69.3	88	31	US-60-197-873-25917	Sequence 25917, A
30	411	65.8	222	1	PCT-US01-08656-6582	Sequence 6582, Ap
31	400	64.0	98	23	US-09-621-976-7034	Sequence 7034, Ap
32	400	64.0	98	20	US-09-834-366-20032	Sequence 20032, A
33	400	64.0	98	31	US-60-147-499-7034	Sequence 7034, Ap
34	400	64.0	98	31	US-60-197-873-20032	Sequence 20032, A
35	381.5	61.0	130	22	US-09-791-537-6190	Sequence 6190, Ap
36	365	58.4	109	28	US-10-221-276-11949	Sequence 11949, Ap
37	359	57.4	147	22	US-09-786-797B-10	Sequence 10, Appl 1
38	359	57.4	147	31	US-60-131-321-8	Sequence 8, Appl 1
39	359	57.4	147	31	US-60-443-566-4508	Sequence 4508, Ap
40	359	57.4	147	31	US-60-455-441-8178	Sequence 8178, Ap
41	359	57.4	147	31	US-60-465-241-8178	Sequence 8178, Ap
42	349	55.8	167	22	US-09-758-472-7597	Sequence 7597, Ap
43	349	55.8	167	28	US-10-235-926-7597	Sequence 7597, Ap
44	317	50.7	82	22	US-09-758-450-564	Sequence 564, Ap
45	317	50.7	82	28	US-10-227-582-564	Sequence 564, App

ALIGNMENTS

RESULT 1  
US-09-904-603-1  
Sequence 1, Application US/09904603  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
Goli, Surya K.  
TITLE OF INVENTION: NOVEL MICROTUBULE-ASSOCIATED PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Inocyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/904,603  
FILING DATE: 12-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

```

1      APPLICATION NUMBER: 08/805,117
2      FILING DATE: <Unknown>
3      ATTORNEY/AGENT INFORMATION:
4      NAME: Billings, Lucy J.
5      REGISTRATION NUMBER: 36,749
6      REFERENCE/DOCKET NUMBER: PF-0211 US
7      TELECOMMUNICATION INFORMATION:
8      TELEPHONE: 415-855-0555
9      TELEFAX: 415-845-4166
10     TELEX: <Unknown>
11
12     INFORMATION FOR SEQ ID NO: 1:
13     SEQUENCE CHARACTERISTICS:
14         .LENGTH: 121 amino acids
15         .TYPE: amino acid
16         .STRANDEDNESS: single
17         .TOPOLOGY: linear
18     IMMEDIATE SOURCE:
19         LIBRARY: THYNOT03
20         CLONE: 144378
21
22     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
23
24     --
25     JS-09-904-603-1

```



```
FEATURE:
NAME/KEY: UNSURE
LOCATION: 105
OTHER INFORMATION: Xaa - Glu,Gln
US-60-197-873-14523

Query Match
Best Local Similarity 87.2%; Score 545; DB 31; Length 107;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPSPRPKORRSFADRCKEVQOIRDOHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
DB 1 MPSPRPKORRSFADRCKEVQOIRDOHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60

QY 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPIADIYEQEKDEDFLYMVYASQETFG 107
DB 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPIADIYEQEKDXG 107

RESULT 9
US-09-834-366-15699
Sequence 15699, Application US/09834366
GENERAL INFORMATION:
APPLICANT: Bejanin, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Jobert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81.US2.REG
CURRENT APPLICATION NUMBER: US/09/834,366
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/197,873
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: Patent.pm
SEQ ID NO 15699
LENGTH: 123
TYPE: PRT
ORGANISM: Homo sapiens
US-09-834-366-15699

Query Match
Best Local Similarity 83.4%; Score 521; DB 23; Length 123;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSPRPKORRSFADRCKEVQOIRDOHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
DB 1 MPSEKTKORRTEQREVEDVRLIREQHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60

QY 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPIADIYEQEKDEDFLYMVYASQETFG 120
DB 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPISEVSEKDEDFLYMVYASQETFG 120

RESULT 10
US-60-197-873-15699
Sequence 15699, Application US/60197873
GENERAL INFORMATION:
APPLICANT: Bejanin, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Jobert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81.US1.PRO
CURRENT APPLICATION NUMBER: US/60/197,873
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: Patent.pm
SEQ ID NO 15699
LENGTH: 123
TYPE: PRT
```

```
ORGANISM: Homo sapiens
US-60-197-873-15699

Query Match
Best Local Similarity 83.4%; Score 521; DB 31; Length 123;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSPRPKORRSFADRCKEVQOIRDOHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
DB 1 MPSEKTKORRTEQREVEDVRLIREQHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60

QY 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPIADIYEQEKDEDFLYMVYASQETFG 120
DB 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPISEVSEKDEDFLYMVYASQETFG 120

RESULT 11
US-09-488-725A-2805
Sequence 2805, Application US/09488725A
GENERAL INFORMATION:
APPLICANT: Hyseq Inc
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
FILE REFERENCE: 784.FL.PCT
CURRENT APPLICATION NUMBER: US/09/488,725A
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US/09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US09/653,450
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US09/662,191
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: US09/693,036
PRIOR FILING DATE: 2000-10-19
PRIOR APPLICATION NUMBER: US09/727,344
PRIOR FILING DATE: 2000-11-29
NUMBER OF SEQ ID NOS: 7144
SOFTWARE: PL_FL_genes_b Versions 1.0
SEQ ID NO 2805
LENGTH: 126
TYPE: PRT
ORGANISM: Homo sapiens
US-09-488-725A-2805

Query Match
Best Local Similarity 83.4%; Score 521; DB 18; Length 125;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSPRPKORRSFADRCKEVQOIRDOHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
DB 1 MPSEKTKORRTEQREVEDVRLIREQHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60

QY 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPIADIYEQEKDEDFLYMVYASQETFG 120
DB 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPISEVSEKDEDFLYMVYASQETFG 120

RESULT 12
US-09-834-366-20036
Sequence 20036, Application US/09834366
GENERAL INFORMATION:
APPLICANT: Bejanin, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Jobert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81.US2.REG
```

```

: CURRENT APPLICATION NUMBER: US/09/834,366
: CURRENT FILING DATE: 2001-04-13
: PRIOR APPLICATION NUMBER: US 60/197,873
: PRIOR FILING DATE: 2000-04-18
: NUMBER OF SEQ ID NOS: 52153
: SOFTWARE: Patent.pm
: SEQ ID NO: 20036
: LENGTH: 125
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-834-366-20036

Query Match
: 83.4%: Score 521; DB 23; Length 125;
: Best Local Similarity 82.5%; Pred. No. 5,1e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSDRPFKORSFADRCKEVOQIRDOHPSKIPVITIERKKGKQLPVLTKIKFLVDDHYNM 60
DB 1 MPSEKTFKQRTFEQVDFVRLIRQHPKIPVITIERKKGKQLPVLTKIKFLVDDHYNM 60
QY 61 SELVKIIRRLQNLNFAFLVNOHSMVSVSTPIADYIPEQKEDGFLYMWYASQETFG 120
DB 61 SELIKIIRRLQNLNANQAFLLVNOHSMVSVSTPISEYSEKEDGFLYMWYASQETFG 120

RESULT 13
US-10-219-051B-7991
: Sequence 7591, Application US/10219051B
: GENERAL INFORMATION:
: APPLICANT: The General Hospital Corporation doing business as Massachusetts General
: APPLICANT: Hospital / Bayer AG
: TITLE OF INVENTION: Nucleotide sequences involved in pain
: FILE REFERENCE: Lea 35693 Foreign Countries
: CURRENT APPLICATION NUMBER: US/10/219,051B
: CURRENT FILING DATE: 2003-05-09
: PRIOR APPLICATION NUMBER: US 60/312,147
: PRIOR FILING DATE: 2001-08-14
: PRIOR APPLICATION NUMBER: US 60/346,382
: PRIOR FILING DATE: 2001-11-01
: PRIOR APPLICATION NUMBER: US 60/333,347
: PRIOR FILING DATE: 2001-11-26
: NUMBER OF SEQ ID NOS: 14715
: SOFTWARE: Perl script
: SEQ ID NO: 7991
: LENGTH: 125
: TYPE: PRT
: ORGANISM: Homo sapiens
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: Refseq / NP_073729
: DATABASE ENTRY DATE: 2002-11-05
US-10-219-051B-7991

Query Match
: 83.4%: Score 521; DB 28; Length 125;
: Best Local Similarity 82.5%; Pred. No. 5,1e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSDRPFKORSFADRCKEVOQIRDOHPSKIPVITIERKKGKQLPVLTKIKFLVDDHYNM 60
DB 1 MPSEKTFKQRTFEQVDFVRLIRQHPKIPVITIERKKGKQLPVLTKIKFLVDDHYNM 60
QY 61 SELVKIIRRLQNLNFAFLVNOHSMVSVSTPIADYIPEQKEDGFLYMWYASQETFG 120
DB 61 SELIKIIRRLQNLNANQAFLLVNOHSMVSVSTPISEYSEKEDGFLYMWYASQETFG 120

RESULT 14
US-10-219-051B-13087
: Sequence 13087, Application US/10219051B
: GENERAL INFORMATION:
: APPLICANT: The General Hospital Corporation doing business as Massachusetts General
: APPLICANT: Hospital / Bayer AG
: TITLE OF INVENTION: Nucleotide sequences involved in pain
: FILE REFERENCE: Lea 35693 Foreign Countries
```

```

: CURRENT APPLICATION NUMBER: US/10/219,051B
: CURRENT FILING DATE: 2003-05-09
: PRIOR APPLICATION NUMBER: US 60/312,147
: PRIOR FILING DATE: 2001-08-14
: PRIOR APPLICATION NUMBER: US 60/346,382
: PRIOR FILING DATE: 2001-11-01
: PRIOR APPLICATION NUMBER: US 60/333,347
: PRIOR FILING DATE: 2001-11-26
: NUMBER OF SEQ ID NOS: 14715
: SOFTWARE: Perl script
: SEQ ID NO: 13087
: LENGTH: 125
: TYPE: PRT
: ORGANISM: Homo sapiens
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: Refseq / NP_073729
: DATABASE ENTRY DATE: 2002-11-05
US-10-219-051B-13087

Query Match
: 83.4%: Score 521; DB 28; Length 125;
: Best Local Similarity 82.5%; Pred. No. 5,1e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSDRPFKORSFADRCKEVOQIRDOHPSKIPVITIERKKGKQLPVLTKIKFLVDDHYNM 60
DB 1 MPSEKTFKQRTFEQVDFVRLIRQHPKIPVITIERKKGKQLPVLTKIKFLVDDHYNM 60
QY 61 SELVKIIRRLQNLNFAFLVNOHSMVSVSTPIADYIPEQKEDGFLYMWYASQETFG 120
DB 61 SELIKIIRRLQNLNANQAFLLVNOHSMVSVSTPISEYSEKEDGFLYMWYASQETFG 120

RESULT 15
US-60-197-873-20036
: Sequence 20036, Application US/60197873
: GENERAL INFORMATION:
: APPLICANT: Bejanin, Stephanie
: APPLICANT: Tanaka, Hiroaki
: APPLICANT: Dumas Milne Edwards, Jean Baptiste
: APPLICANT: Joebert, Severin
: APPLICANT: Giordano, Jean-Yves
: TITLE OF INVENTION: ESTs and Encoded Human Proteins.
: FILE REFERENCE: 81,051,PRO
: CURRENT APPLICATION NUMBER: US/60/197,873
: CURRENT FILING DATE: 2000-04-18
: NUMBER OF SEQ ID NOS: 52153
: SOFTWARE: Patent.pm
: SEQ ID NO: 20036
: LENGTH: 125
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-197-873-20036

Query Match
: 83.4%: Score 521; DB 31; Length 125;
: Best Local Similarity 82.5%; Pred. No. 5,1e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSDRPFKORSFADRCKEVOQIRDOHPSKIPVITIERKKGKQLPVLTKIKFLVDDHYNM 60
DB 1 MPSEKTFKQRTFEQVDFVRLIRQHPKIPVITIERKKGKQLPVLTKIKFLVDDHYNM 60
QY 61 SELVKIIRRLQNLNFAFLVNOHSMVSVSTPIADYIPEQKEDGFLYMWYASQETFG 120
DB 61 SELIKIIRRLQNLNANQAFLLVNOHSMVSVSTPISEYSEKEDGFLYMWYASQETFG 120

Search completed: August 11, 2003, 10:23:10
Job time : 375 secs
```

**THIS PAGE BLANK (USPTO)**



```
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq Inc
;; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;; FILE REFERENCE: 784FLPCT
;; CURRENT APPLICATION NUMBER: US/10/286,897
;; CURRENT FILING DATE: 2002-11-01
;; PRIOR APPLICATION NUMBER: US/09/488,725
;; PRIOR FILING DATE: 2000-01-21
;; PRIOR APPLICATION NUMBER: US/09/552,317
;; PRIOR FILING DATE: 2000-04-25
;; PRIOR APPLICATION NUMBER: US/09/598,042
;; PRIOR FILING DATE: 2000-06-20
;; PRIOR APPLICATION NUMBER: US/09/620,312
;; PRIOR FILING DATE: 2000-07-19
;; PRIOR APPLICATION NUMBER: US/09/653,450
;; PRIOR FILING DATE: 2000-08-31
;; PRIOR APPLICATION NUMBER: US/09/662,191
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: US/09/693,036
;; PRIOR FILING DATE: 2000-10-19
;; PRIOR APPLICATION NUMBER: US/09/727,344
;; PRIOR FILING DATE: 2000-11-29
;; NUMBER OF SEQ ID NOS: 7143
;; SOFTWARE: pf-fl_genes_b Versions 1.0
;; SEQ ID NO: 6377
;; LENGTH: 155
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-286-897-6377

Query Match
Best Local Similarity 83.4%; Score 521; DB 6; Length 155;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSPDRPFKORSFADCKEVOQIRDOHPSKIPVIERKGEKQLPVLDKTKFLVPDHYNM 60
    |||:|||||:|:|:|:|||||:|||||:|||||:|||||:|||||:|||||:
DB 31 MPSEKTFKQRTFEQVRVDVRLIRQHPKIPVIERKGEKQLPVLDKTKFLVPDHYNM 90

QY 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPIADIYEOEKDEGFLVMYVASOETFG 120
    |||:|||||:|:|:|:|||||:|||||:|||||:|||||:|||||:
DB 91 SELKIIIRRLQLNANOAFLLVNGHSMVSVSTPISEYSEKEDGFLVMYVASOETFG 150

RESULT 3
US-10-273-573-6982
;; Sequence 6982, Application US/10273573
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc
;; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
;; FILE REFERENCE: 21272-066
;; CURRENT APPLICATION NUMBER: US/10/273,573
;; CURRENT FILING DATE: 2002-10-18
;; PRIOR APPLICATION NUMBER: 09/522,929
;; PRIOR FILING DATE: 2000-04-18
;; PRIOR APPLICATION NUMBER: 09/770,160
;; PRIOR FILING DATE: 2001-01-26
;; NUMBER OF SEQ ID NOS: 10994
;; SOFTWARE: Custom
;; SEQ ID NO: 6982
;; LENGTH: 222
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: DOMAIN
;; LOCATION: (9)...(31)
;; OTHER INFORMATION: PROSTANOID EPI RECEPTOR SIGNATURE domain identified by
;; OTHER INFORMATION: EMATRIX, accession number PR00580E, p-value=2.014e-10, raw score
;; OTHER INFORMATION: 7.15
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(222)
;; OTHER INFORMATION: Xaa = X or * as defined in Table 2.
US-10-273-573-6982
```

```
Query Match
Best Local Similarity 65.8%; Score 411; DB 6; Length 222;
Matches 82; Conservative 16; Mismatches 22; Indels 2; Gaps 2;

QY 1 MSPDRPFKORSFADCKEVOQIRDOHPSKIPVIERKGEKQLPVLDKTKFLVPDHYNM 59
    |||:|||||:|:|:|:|||||:|||||:|||||:|||||:|||||:
DB 32 MPSEKTFKQRTFEQVRVDVRLIRQHPKIPVIERKGEKQLPVLDKTKFLVPDHYNM 91

QY 60 MSELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPIADIYEOEKDEGFLVMYVASOET 118
    |||:|||||:|:|:|:|||||:|||||:|||||:|||||:|||||:
DB 92 MSELKIIIRRLQLNANOAFLLVNGHSMVSVSTPISEYSEKEDGFLVMYVASOET 151

QY 119 FG 120
    :|
DB 152 YG 153

RESULT 4
US-10-293-244-3825
;; Sequence 3825, Application US/10293244
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;; FILE REFERENCE: 21272-029
;; CURRENT APPLICATION NUMBER: US/10/293,244
;; CURRENT FILING DATE: 2002-11-12
;; PRIOR APPLICATION NUMBER: Not Yet Assigned
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: 09/728,422
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: 09/693,325
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 09/663,561
;; PRIOR FILING DATE: 2000-09-15
;; PRIOR APPLICATION NUMBER: 09/654,936
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 09/620,325
;; PRIOR FILING DATE: 2000-07-19
;; PRIOR APPLICATION NUMBER: 09/598,075
;; PRIOR FILING DATE: 2000-06-20
;; PRIOR APPLICATION NUMBER: 09/560,875
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 09/496,914
;; PRIOR FILING DATE: 2000-02-03
;; NUMBER OF SEQ ID NOS: 3960
;; SOFTWARE: Custom
;; SEQ ID NO: 3825
;; LENGTH: 120
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-293-244-3825

Query Match
Best Local Similarity 48.2%; Score 301; DB 6; Length 120;
Matches 58; Conservative 16; Mismatches 24; Indels 0; Gaps 0;

QY 23 IRDQHPKIPVIERKGEKQLPVLDKTKFLVPDHYNMSELVKIIRRLQLNPTQAFLL 82
    ||:|||||:|:|:|:|||||:|||||:|||||:|||||:|||||:
DB 2 IRAKFPNKIPVIERKGEKQLPVLDKTKFLVPDHYNMSELVKIIRRLQLNPTQAFLL 61

QY 83 VNQHSMSVSTPIADIYEOEKDEGFLVMYVASOETFG 120
    ||:||||:|:|:|:|||||:|||||:|||||:|||||:|||||:
DB 62 VNKSLSVMSATMAEIRDRYKDEGFLVMYVASOETFG 99

RESULT 5
US-10-293-244-1857
;; Sequence 1857, Application US/10293244
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; APPLICANT: Tang, Y. Tom et al
```



```

Query Match      43.5%  Score 272:  DB 6:  Length 135:
Best local Similarity 55.6%:  Pred. No. 1, 7e-23:
Matches 55:  Conservative 18:  Mismatches 26:  Indels 0:  Gaps
QY      1 MPSDPRKQKRSRADRCREYQOQIRDRHPKSLPIYIRIRYGEKOLPYLDKTKFLVPHVM 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      7 IFSVPRPKQKSLATKQEEVAGIRAFPNKPIPVVYRVPYRPETFLPLDKTKFLVPELTM 66
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      61 SELVKIIRRRLLNPTQAEFLVNVHSMWSVSTPIADIV 99
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      67 TQFLSIIRSRMVRATEAFYLLVNNKSLVMSMTAEI 105

```

Query Match	41.0%	Score	256;	DB	5;	Length	50;
Best Local Similarity	100.0%	Pred. NO.	2.8e-22;				
Matches	50;	Conservative	0;	Mismatches	0;	Indels	0;

APPLICANT: RUDEN ET AL.  
TITLE OF INVENTION: 83 Human Secreted Protein  
FILE REFERENCE: P2012P1  
CURRENT APPLICATION NUMBER: US/09/8,201,644  
CURRENT FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: US/09/7,236,557  
PRIOR FILING DATE: 1999-01-26  
PRIOR APPLICATION NUMBER: PCT/US98/15949

```

Query Match      37.8%;  Score 2.6;  DB 5;  length 46;
Best local Similarity 100.0%;  Freq. No. 4;  ye-20;
Matches 46;  Conservative 0;  Mismatches 0;  Indels 0;

76  TAAFLLVNQHSMVSVSTFIALVTHQEKDEAPFLVYVASQETPGF 12
      | | | | | | | | | | | | | | | | | | | | | | | |
1  TAAFLLVNQHSMVSVSTFIALVTHQEKDEAPFLVYVASQETPGF 46

```

US-10-408-765A-1021  
 RESULT 8  
 Sequence 1021, Application US/10408765A  
 GENERAL INFORMATION:  
 APPLICANT: Gosh, Sonmitra S.  
 APPLICANT: Fahy, Eoin D.  
 APPLICANT: Zhang, Bing  
 APPLICANT: Glison, Bradford W.  
 APPLICANT: Taylor, Steven W.



```

0Y      3  SDRPEKQRRSAD-----RCK-----EVQG-----IRDOHPSKIPIYIRRYGCK 42
      4  : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : :
Db      832  NDRDFQDEMTLSDKDLIIKCKKANEYENOIEQLQKMLGLDKKKNISIEEFKPLTKGT 891
0Y      43  QLPVLDTKRLFVPHVNMSELVKIIRRLQINPTQAFLLVNMCHSMVSVSTPIADIEOE 102
      44  : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : :
Db      892  FKPIYENK-----NIERLNKSJLKKIELMQLTFYIQINPIYLTKLFNSIYTKNOT 944
0Y      103  KD-EDGF-----LTMVYASQET 119
      45  : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : :
Db      945  KSGQDLFQSVIQLFPVRDSSITY 967

RESULT 12
US-10-617-320-4624
: Sequence 4624, Application US/10617320
: GENERAL INFORMATION:
: APPLICANT: Lynn A Doucette-Stamm and David Bush
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
: SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
: THERAPEUTICS
NUMBER OF SEQUENCES: 5206
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02154
COMPUTER READABLE FORM:
: MEDIUM TYPE: CD-ROM ISO9660
: COMPUTER: <Unknown>
: OPERATING SYSTEM: <Unknown>
: SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/617,320
: FILING DATE: 10-Jul-2003
Prior Application DATA:
: APPLICATION NUMBER: US/09/107,433
: FILING DATE: 30-Jun-1998
: APPLICATION NUMBER: 60/085131
: FILING DATE: May 12, 1998
: APPLICATION NUMBER: 60/051553
: FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
: NAME: Atinello, Pamela Deneke
: REGISTRATION NUMBER: 40,489
: REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
: TELEPHONE: (781)893-5007
: TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4624:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 243 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
MOLECULE TYPE: protein
: HYPOTHEICAL: YES
ORIGINAL SOURCE:
: ORGANISM: Streptococcus pneumoniae
FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (8) LOCATION 1...243
: SEQUENCE DESCRIPTION: SEQ ID NO: 4624:
US-10-617-320-4624

Query Match      10.94; Score 68; DB 6; Length 243;
Best local Similarity 27.7%; Pred. No. 2.4;
Matches 28; Conservative 18; Mismatches 39; Indels 16; Gaps 5;

0Y      8  KURSFPADRCQEVQOIRDOHPSKIPIVIT--ERYKGEQLPVLDKTKFLVPHVNMSELVK 65
      9  : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : : 11 : :

```

```

Db      151  KKKEEFAVNHITITLKEQYKIELMMILVKSLSLIGROOLRLQD---LFRKAQVWJGVII 204
QY      66  IIRRLQINPTQAFILVNOHSMVSVSLPIALITEQEKD 106
       205  LFLATLEIKTQF-LIIIVQFES-----FDIYILMERKFE 247

RESULT 13
US-09-820-843B-16
: Sequence 16, Application US/098208438
: GENERAL INFORMATION:
: APPLICANT: Council of Scientific and Industrial Research
: TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION of CANDIDATE
: FILE REFERENCE: Q63915
: CURRENT APPLICATION NUMBER: US/09/820-8438
: CURRENT FILING DATE: 2001-03-10
: NUMBER OF SEQ ID NOS: 118
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 16
: LENGTH: 1616
: TYPE: PRT
: ORGANISM: M. genitalium
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: cytochrome-oxidase accessory protein
: NAME/KEY: misc_feature
: OTHER INFORMATION: gii1046097
US-09-820-843B-16

Query Match      10.9%; Score 68; DB 5; Length 1616;
Best Local Similarity 20.1%; Pred. No. 29;
Matches 31; Conservative 25; Mismatches 40; Indels 58; Gaps 9;

QY      4  DRPF-----KGRSEFAIDCKEVOGIRQHSKIPVIERKKGKQLVILKTK--- 51
       45  DQPLDGLKKNILKTKTSFNSKRKEIVYKSKSP-----IUFDEIKRGV 89
Db      52  FLVDP-----IVNMSEIVKILRRKLQINPTQAFILVNOHSMVSVSLPIALITEQEK 102
       90  FIVPETILSKRLNOKTVNTVNPJNSVPT-----INENKLVLEINNPEIKVLFTK 144
QY      103  KD-----ED--GFLVWVVASQELHPF 121
       144  KDSFTTIKREKLNEDSQAFWY-IFWJDKKYCF 176
Db

RESULT 14
US-10-603-113-15278
: Sequence 15278, Application US/10603113
: GENERAL INFORMATION:
: APPLICANT: Keith Welislock et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICI
: FILE REFERENCE: 107196.132
: CURRENT APPLICATION NUMBER: US/10/603-113
: CURRENT FILING DATE: 2003-06-24
: PRIOR APPLICATION NUMBER: US/09/248,796
: NUMBER OF SEQ ID NOS: 28206
: SEQ ID NO 15278
: LENGTH: 469
: TYPE: PRT
: ORGANISM: Candida albicans
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (7)-(14)
: OTHER INFORMATION: Identity of amino acid sequences at the above locations are un
US-10-603-113-15278

Query Match      10.6%; Score 66.5; DB 6; Length 469;
Best Local Similarity 25.2%; Pred. No 8.4;

```

Matches 31; Conservative 22; Mismatches 39; Indels 31; Gaps 8;

```
QY      16 RCKEVOOI-RDQHSKIPVIERKYGK-----QLPVLDKTKFLV---POHVNMSLVKI 66
      | : : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      159 RSHQVEKLEFKENKINIPILISSPSNGEKPPETDLP--DQLEDLIVASPDYINCTTIIIT 216
      | : : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY      67 IRRR---LQLNPTQAEFF---LLVNOHSMVSVSTPIADIYEQEKDEDEGFLYMWYASQET 118
      | : : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      217 IHDPPEIWANINPTNGVLETNSTWVLEHPLVQ---IMDKIQEKEDPE-----NQEK 264
      | : : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY      119 FGF 121
      | : : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      265 Fgy 267
```

## RESULT 15

```
US-10-603-113-17078
; Sequence 17078, Application US/10603113
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/10/603,113
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/248,796
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO 17078
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-603-113-17078
```

Query Match 10.6%; Score 66; DB 6; Length 454;

Best Local Similarity 22.7%; Pred. No. 9.1;

Matches 25; Conservative 21; Mismatches 54; Indels 10; Gaps 1;

```
QY      8 KORRSFADRCKEVOQIRDOHPSKIPVIERKYGKQLPVLDKTKFLVDPDHVNMSELVKII 67
      | : : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      233 KYRKFASRLK-----NKLSTRYRPGVGESYLDVLTRLRLPLIAEVERTIDHLLII 282
      | : : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY      68 RRRLQNLPTQAEFLLVNOHSMVSVSTPIADIYEQEKDEDEGFLYMWYASQE 117
      | : : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      283 SHRVSRLILAYFLINDKSAIGELDVPILHTLYCLELHPYGTGYTMEYDE 332
```

Search completed: August 11, 2003, 10:23:37  
Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

us nucleic - nucleic search, using sw model

Run on: August 15, 2003, 08:09:17 : Search time 2934 Seconds  
(without alignments)  
7147.724 Million cell updates/sec

Title: US-09-904-603-2  
Perfect score: 640  
Sequence: 1 CTCGCCGACGCGCGCGCC.....GCCGCCCTAGTCATACAGCCCA 640

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32997241 seqs, 16383922548 residues

Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/prodata/1/pna/PC705\_COMB.seq.\*  
2: /cgn2\_6/prodata/1/pna/PC705\_COMB.seq.OLD.\*  
3: /cgn2\_6/prodata/1/pna/US06\_COMB.seq.\*  
4: /cgn2\_6/prodata/1/pna/US07\_COMB.seq.\*  
5: /cgn2\_6/prodata/1/pna/US080\_COMB.seq.\*  
6: /cgn2\_6/prodata/1/pna/US081\_COMB.seq.\*  
7: /cgn2\_6/prodata/1/pna/US082\_COMB.seq.\*  
8: /cgn2\_6/prodata/1/pna/US083\_COMB.seq.\*  
9: /cgn2\_6/prodata/1/pna/US084\_COMB.seq.\*  
10: /cgn2\_6/prodata/1/pna/US085\_COMB.seq.\*  
11: /cgn2\_6/prodata/1/pna/US086\_COMB.seq.\*  
12: /cgn2\_6/prodata/1/pna/US087\_COMB.seq.\*  
13: /cgn2\_6/prodata/1/pna/US088\_COMB.seq.\*  
14: /cgn2\_6/prodata/1/pna/US089\_COMB.seq.\*  
15: /cgn2\_6/prodata/1/pna/US090\_COMB.seq.\*  
16: /cgn2\_6/prodata/1/pna/US091\_COMB.seq.\*  
17: /cgn2\_6/prodata/1/pna/US092A\_COMB.seq.\*  
18: /cgn2\_6/prodata/1/pna/US092B\_COMB.seq.\*  
19: /cgn2\_6/prodata/1/pna/US093A\_COMB.seq.\*  
20: /cgn2\_6/prodata/1/pna/US093B\_COMB.seq.\*  
21: /cgn2\_6/prodata/1/pna/US094\_COMB.seq.\*  
22: /cgn2\_6/prodata/1/pna/US095A\_COMB.seq.\*  
23: /cgn2\_6/prodata/1/pna/US095B\_COMB.seq.\*  
24: /cgn2\_6/prodata/1/pna/US095C\_COMB.seq.\*  
25: /cgn2\_6/prodata/1/pna/US095D\_COMB.seq.\*  
26: /cgn2\_6/prodata/1/pna/US096A\_COMB.seq.\*  
27: /cgn2\_6/prodata/1/pna/US096B\_COMB.seq.\*  
28: /cgn2\_6/prodata/1/pna/US096C\_COMB.seq.\*  
29: /cgn2\_6/prodata/1/pna/US096D\_COMB.seq.\*  
30: /cgn2\_6/prodata/1/pna/US096E\_COMB.seq.\*  
31: /cgn2\_6/prodata/1/pna/US097A\_COMB.seq.\*  
32: /cgn2\_6/prodata/1/pna/US097B\_COMB.seq.\*  
33: /cgn2\_6/prodata/1/pna/US097C\_COMB.seq.\*  
34: /cgn2\_6/prodata/1/pna/US098A\_COMB.seq.\*  
35: /cgn2\_6/prodata/1/pna/US098B\_COMB.seq.\*  
36: /cgn2\_6/prodata/1/pna/US098C\_COMB.seq.\*  
37: /cgn2\_6/prodata/1/pna/US098D\_COMB.seq.\*  
38: /cgn2\_6/prodata/1/pna/US099A\_COMB.seq.\*  
39: /cgn2\_6/prodata/1/pna/US099B\_COMB.seq.\*  
40: /cgn2\_6/prodata/1/pna/US099C\_COMB.seq.\*  
41: /cgn2\_6/prodata/1/pna/US099D\_COMB.seq.\*  
42: /cgn2\_6/prodata/1/pna/US099E\_COMB.seq.\*  
43: /cgn2\_6/prodata/1/pna/US099F\_COMB.seq.\*

44: /cgn2\_6/prodata/1/pna/US100A\_COMB.seq.\*  
45: /cgn2\_6/prodata/1/pna/US100B\_COMB.seq.\*  
46: /cgn2\_6/prodata/1/pna/US101A\_COMB.seq.\*  
47: /cgn2\_6/prodata/1/pna/US101B\_COMB.seq.\*  
48: /cgn2\_6/prodata/1/pna/US102A\_COMB.seq.\*  
49: /cgn2\_6/prodata/1/pna/US102B\_COMB.seq.\*  
50: /cgn2\_6/prodata/1/pna/US103A\_COMB.seq.\*  
51: /cgn2\_6/prodata/1/pna/US103B\_COMB.seq.\*  
52: /cgn2\_6/prodata/1/pna/US104A\_COMB.seq.\*  
53: /cgn2\_6/prodata/1/pna/US104B\_COMB.seq.\*  
54: /cgn2\_6/prodata/1/pna/US6000\_COMB.seq.\*  
55: /cgn2\_6/prodata/1/pna/US6001\_COMB.seq.\*  
56: /cgn2\_6/prodata/1/pna/US6002\_COMB.seq.\*  
57: /cgn2\_6/prodata/1/pna/US6003\_COMB.seq.\*  
58: /cgn2\_6/prodata/1/pna/US6004\_COMB.seq.\*  
59: /cgn2\_6/prodata/1/pna/US6005\_COMB.seq.\*  
60: /cgn2\_6/prodata/1/pna/US6006\_COMB.seq.\*  
61: /cgn2\_6/prodata/1/pna/US6007\_COMB.seq.\*  
62: /cgn2\_6/prodata/1/pna/US6008\_COMB.seq.\*  
63: /cgn2\_6/prodata/1/pna/US6009\_COMB.seq.\*  
64: /cgn2\_6/prodata/1/pna/US6010\_COMB.seq.\*  
65: /cgn2\_6/prodata/1/pna/US6011\_COMB.seq.\*  
66: /cgn2\_6/prodata/1/pna/US6012\_COMB.seq.\*  
67: /cgn2\_6/prodata/1/pna/US6013\_COMB.seq.\*  
68: /cgn2\_6/prodata/1/pna/US6014\_COMB.seq.\*  
69: /cgn2\_6/prodata/1/pna/US6015\_COMB.seq.\*  
70: /cgn2\_6/prodata/1/pna/US6016\_COMB.seq.\*  
71: /cgn2\_6/prodata/1/pna/US6017\_COMB.seq.\*  
72: /cgn2\_6/prodata/1/pna/US6018\_COMB.seq.\*  
73: /cgn2\_6/prodata/1/pna/US6019\_COMB.seq.\*  
74: /cgn2\_6/prodata/1/pna/US6020\_COMB.seq.\*  
75: /cgn2\_6/prodata/1/pna/US6021\_COMB.seq.\*  
76: /cgn2\_6/prodata/1/pna/US6022\_COMB.seq.\*  
77: /cgn2\_6/prodata/1/pna/US6023A\_COMB.seq.\*  
78: /cgn2\_6/prodata/1/pna/US6023B\_COMB.seq.\*  
79: /cgn2\_6/prodata/1/pna/US6024\_COMB.seq.\*  
80: /cgn2\_6/prodata/1/pna/US6025\_COMB.seq.\*  
81: /cgn2\_6/prodata/1/pna/US6026\_COMB.seq.\*  
82: /cgn2\_6/prodata/1/pna/US6027\_COMB.seq.\*  
83: /cgn2\_6/prodata/1/pna/US6028\_COMB.seq.\*  
84: /cgn2\_6/prodata/1/pna/US6029\_COMB.seq.\*  
85: /cgn2\_6/prodata/1/pna/US6030\_COMB.seq.\*  
86: /cgn2\_6/prodata/1/pna/US6031\_COMB.seq.\*  
87: /cgn2\_6/prodata/1/pna/US6032\_COMB.seq.\*  
88: /cgn2\_6/prodata/1/pna/US6033\_COMB.seq.\*  
89: /cgn2\_6/prodata/1/pna/US6034\_COMB.seq.\*  
90: /cgn2\_6/prodata/1/pna/US6035\_COMB.seq.\*  
91: /cgn2\_6/prodata/1/pna/US6036\_COMB.seq.\*  
92: /cgn2\_6/prodata/1/pna/US6037\_COMB.seq.\*  
93: /cgn2\_6/prodata/1/pna/US6038\_COMB.seq.\*  
94: /cgn2\_6/prodata/1/pna/US6039\_COMB.seq.\*  
95: /cgn2\_6/prodata/1/pna/US6040\_COMB.seq.\*  
96: /cgn2\_6/prodata/1/pna/US6041\_COMB.seq.\*  
97: /cgn2\_6/prodata/1/pna/US6042\_COMB.seq.\*  
98: /cgn2\_6/prodata/1/pna/US6043\_COMB.seq.\*  
99: /cgn2\_6/prodata/1/pna/US6044\_COMB.seq.\*  
100: /cgn2\_6/prodata/1/pna/US6045\_COMB.seq.\*  
101: /cgn2\_6/prodata/1/pna/US6046\_COMB.seq.\*  
102: /cgn2\_6/prodata/1/pna/US6047\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length (bp)	%	Description
1	640	640	100.0	US-09-904-603-2
2	627.4	925	25	US-09-572-411-5453
3	627.4	945	47	US-10-170-245-27458
4	627.4	977	21	US-09-471-275-1555

5 627.4 98.0 977 21 US-09-488-725B-8516 Sequence 8516, App  
6 627.4 98.0 1000 1 PCT-US01-18569-361 Sequence 361, App  
7 627.4 98.0 1000 2 PCT-US01-18569-361 Sequence 361, App  
8 627.4 98.0 1000 49 US-10-264-049-361 Sequence 361, App  
9 627.4 98.0 1000 77 US-60-234-690-7194 Sequence 7194, App  
10 627.4 98.0 1009 87 US-60-323-843-7194 Sequence 7194, App  
11 627.4 98.0 1019 52 US-10-405-027-30 Sequence 30, App  
12 627.4 98.0 1356 21 US-09-428-151A-9451 Sequence 9451, App  
13 627 98.0 951 1 PCT-US98-15949-94 Sequence 94, App  
14 627 98.0 951 1 PCT-US98-15949-94 Sequence 94, App  
15 605.4 94.6 1017 71 US-60-172-373-10400 Sequence 10400, A  
16 593 92.7 719 28 US-09-644-871-273 Sequence 273, App  
17 507 79.2 539 20 US-09-354-899-5690 Sequence 5690, App  
18 507 79.2 539 21 US-09-428-151A-5852 Sequence 5852, App  
19 497 77.7 557 20 US-09-397-761A-2899 Sequence 2899, App  
20 497 77.7 557 35 US-09-838-601-2899 Sequence 2899, App  
21 493 77.0 505 20 US-09-371-168-264 Sequence 264, App  
22 493 77.0 505 34 US-09-801-833-264 Sequence 264, App  
23 482.2 75.3 605 80 US-60-253-652-16982 Sequence 16982, A  
24 480.4 75.1 600 82 US-60-253-652-14610 Sequence 14610, A  
25 475.2 72.3 570 28 US-09-652-109-5783 Sequence 5783, App  
26 462.6 72.3 557 23 US-10-405-027-1218 Sequence 1218, App  
27 461.2 72.1 557 47 US-09-534-854-14761 Sequence 14761, A  
28 460.4 71.9 966 82 US-10-170-235-26846 Sequence 26846, A  
29 460.4 71.9 966 82 US-09-278-258-12073 Sequence 12073, A  
30 460.4 71.9 1014 17 US-09-236-557-26 Sequence 26, App  
31 460.4 71.9 1014 47 US-10-160-162-26 Sequence 26, App  
32 456.4 71.3 704 1 PCT-US01-08631-9784 Sequence 9784, App  
33 456.4 71.3 704 2 PCT-US01-08631-9784 Sequence 9784, App  
34 455.4 71.2 516 28 US-09-644-871-3509 Sequence 3509, App  
35 452.4 70.7 792 17 US-09-205-070-14460 Sequence 14460, A  
36 452.4 70.7 792 19 US-09-340-623-14460 Sequence 14460, A  
37 452.4 70.7 792 20 US-09-359-922-867 Sequence 867, App  
38 452.4 70.7 792 20 US-09-359-922-867 Sequence 867, App  
39 452.4 70.7 792 37 US-09-898-888-14460 Sequence 14460, A  
40 452.4 70.7 792 37 US-09-898-888-14460 Sequence 14460, A  
41 452.4 70.7 792 38 US-09-919-002-867 Sequence 867, App  
42 450.6 70.4 470 22 US-09-528-409-60388 Sequence 60388, A  
43 450.6 70.4 470 39 US-09-933-524-60388 Sequence 60388, A  
44 450.6 70.4 470 39 US-09-933-524A-60388 Sequence 60388, A  
45 449.8 70.3 516 25 US-09-572-411-682 Sequence 682, App

## ALIGNMENTS

RESULT 1  
US-09-904-603-2  
Sequence 2, Application US/09904603  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
Goli, Surya K.  
TITLE OF INVENTION: NOVEL MICROTUBULE-ASSOCIATED PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Inocyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/904,603  
FILING DATE: 12-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/805,117  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0211 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX: <unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 640 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PHYNOT03  
CLONE: 1441378  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-904-603-2  
Query Match 100.0%; Score 640; DB 38; Length 640;  
Best Local Similarity 100.0%; Pred. No. 1.6e-109; Mismatches 0; Gaps 0;  
Matches 640; Conservative 0; Mismatches 0; Indels 0;  
1 CTCCGCGACGCGACCGCCGCTGCTCAAGCGGAGCCCGGAGCCCTTGAGCGGAGCGC 60  
1 CTCCGCGACGCGACCGCCGCTGCTCAAGCGGAGCCCGGAGCCCTTGAGCGGAGCGC 60  
61 GGACCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
61 GGACCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
121 GCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
121 GCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
181 CGACCGCTGTAGAGGAGTACAGAGATCCGCGACGACCGCGGCGGCGGCGGCGGCGG 240  
181 CGACCGCTGTAGAGGAGTACAGAGATCCGCGACGACCGCGGCGGCGGCGGCGGCGG 240  
241 CATGAGCGGCTACAGAGGAGTACAGAGATCCGCGGCGGCGGCGGCGGCGGCGGCGG 300  
241 CATGAGCGGCTACAGAGGAGTACAGAGATCCGCGGCGGCGGCGGCGGCGGCGGCGG 300  
301 CCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360  
301 CCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360  
361 CCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420  
361 CCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420  
421 CATGAGCGGCTACAGAGGAGTACAGAGATCCGCGGCGGCGGCGGCGGCGGCGGCGG 480  
421 CATGAGCGGCTACAGAGGAGTACAGAGATCCGCGGCGGCGGCGGCGGCGGCGGCGG 480  
481 CTCCGAGGAAACCTTCGCTTCTGAGCCAGAGAGTGGGGGCGGCGGCGGCGGCGGCGG 540  
481 CTCCGAGGAAACCTTCGCTTCTGAGCCAGAGAGTGGGGGCGGCGGCGGCGGCGGCGG 540  
541 GGGCGGCGGCTACAGAGGAGTACAGAGATCCGCGGCGGCGGCGGCGGCGGCGGCGG 600  
541 GGGCGGCGGCTACAGAGGAGTACAGAGATCCGCGGCGGCGGCGGCGGCGGCGGCGG 600  
601 GGTGCGGCTGAGGAGGAGTACAGAGATCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 640  
601 GGTGCGGCTGAGGAGGAGTACAGAGATCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 640

RESULT 2  
US-09-572-411-5453  
Sequence 5453, Application US/09572411  
GENERAL INFORMATION:

```

? APPLICANT: Gearling, David P.
? APPLICANT: Holzman, Douglas A.
? APPLICANT: Xu, Yong Yao
? APPLICANT: Axton, Susan L.
? APPLICANT: Shyjan, Andrew W.
? APPLICANT: Fraser, Christopher C.
? TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
? FILE REFERENCE: THEREFOR
? CURRENT APPLICATION NUMBER: US/09/572,411
? EARLIER FILING DATE: 2000-05-12
? EARLIER APPLICATION NUMBER: 60/133,993
? EARLIER FILING DATE: 1999-05-13
? EARLIER APPLICATION NUMBER: 60/135,617
? EARLIER FILING DATE: 1999-05-24
? EARLIER APPLICATION NUMBER: 60/135,816
? EARLIER FILING DATE: 1999-05-24
? EARLIER APPLICATION NUMBER: 60/135,817
? EARLIER FILING DATE: 1999-05-24
? EARLIER APPLICATION NUMBER: 60/135,950
? EARLIER FILING DATE: 1999-05-24
? EARLIER APPLICATION NUMBER: 60/135,679
? EARLIER FILING DATE: 1999-05-24
? EARLIER APPLICATION NUMBER: 60/135,630
? NUMBER OF SEQ ID NOS: 7417
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO: 5453
? LENGTH: 925
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)...(925)
? OTHER INFORMATION: n = A,T,C or G
US-09-572-411-5453

```

```

Query Match      98.0%; Score 627.4; DB 25; Length 925;
Best Local Similarity 99.7%; Pred. No. 3,4e-107;
Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTCGCGAGGCGGACGCGCGCTGTCTGACGCGGAGCGCGCGGAGCGCTTACGCGCGAGCGCG 60
DB 38 CTCGCGAGGCGGACGCGCGCTGTCTGACGCGGAGCGCGCGGAGCGCTTACGCGCGAGCGCG 97
QY 61 GGAG-CCCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 119
DB 98 GGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 157
QY 120 GCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 158 GCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 217
QY 180 CGGACCGCTGTAAAGAGGTACAGCAGATCCGCGACCGACCGCGCGCGCGCGCGCGCGCG 239
DB 218 CGGACCGCTGTAAAGAGGTACAGCAGATCCGCGACCGACCGCGCGCGCGCGCGCGCGCG 277
QY 240 TCATCGAGCGCTTACAGAGGTAGAGAGCGAGTGGCGCGCTCTGAGCAGAGCGCGCGCGCG 299
DB 278 TCATCGAGCGCTTACAGAGGTAGAGAGCGAGTGGCGCGCTCTGAGCAGAGCGCGCGCGCG 337
QY 300 TCCGCGAGCGCTTACAGAGGTAGAGAGCGAGTGGCGCGCTCTGAGCAGAGCGCGCGCGCG 359
DB 338 TCCGCGAGCGCTTACAGAGGTAGAGAGCGAGTGGCGCGCTCTGAGCAGAGCGCGCGCGCG 397
QY 360 ACCGCGAGCGCGCTTCTCTCTGCTGTGAGACGACGACGACGATGTAGTGTGTGCGCGCG 419
DB 398 ACCGCGAGCGCGCTTCTCTCTGCTGTGAGACGACGACGACGATGTAGTGTGTGCGCGCG 457
QY 420 CCATCGCGAGCATCTACGAGCGAGGAGAGAGAGAGAGCGCGCTTCTCTATATGGCTAGCG 479
DB 458 CCATCGCGAGCATCTACGAGCGAGGAGAGAGAGAGAGCGCGCTTCTCTATATGGCTAGCG 517

```

```

QY 480 CTCGCGAGGAAACCTTGGCTTGTGAGCGAGTACGCGCGCGCGCGCGCGCGCGCGCGCG 539
DB 518 CTCGCGAGGAAACCTTGGCTTGTGAGCGAGTACGCGCGCGCGCGCGCGCGCGCGCGCG 577
QY 540 GGGCGCGCGCTGAGCGCGCTGCGTACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
DB 578 GGGCGCGCGCTGAGCGCGCTGCGTACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 637
QY 600 TGGTGGCGCTGGCGAGCGATGTGGCGCGCTTACGAGAGCGCA 640
DB 638 TGGTGGCGCTGGCGAGCGATGTGGCGCGCTTACGAGAGCGCA 678

RESULT 3
US-10-170-235-27458
? Sequence 27458, Application US/10170245
? GENERAL INFORMATION:
? APPLICANT: VENTER, J. Craig
? TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU
? TITLE OF INVENTION: TRANSCRIPTS, FOR DEPICTING EXPRESSION AND OTHER USES THEREOF
? FILE REFERENCE: C1001380
? CURRENT APPLICATION NUMBER: US/10/170,235
? CURRENT FILING DATE: 2003-03-17
? NUMBER OF SEQ ID NOS: 42514
? SEQ ID NO 27458
? LENGTH: 935
? TYPE: DNA
? ORGANISM: HUMAN
US-10-170-235-27458

```

```

Query Match      98.0%; Score 627.4; DB 47; Length 945;
Best Local Similarity 99.7%; Pred. No. 3,4e-107;
Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTCGCGAGGCGGAGCGCGCGCTGTCTGACGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 67 CTCGCGAGGCGGAGCGCGCGCTGTCTGACGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 126
QY 61 GGAG-CCCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 119
DB 127 GGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 186
QY 120 GCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 187 GCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 246
QY 180 CGGACCGCTGTAAAGAGGTACAGCAGATCCGCGACCGACCGCGCGCGCGCGCGCGCGCG 239
DB 247 CGGACCGCTGTAAAGAGGTACAGCAGATCCGCGACCGACCGCGCGCGCGCGCGCGCGCG 306
QY 240 TCATCGAGCGCTTACAGAGGTAGAGAGCGAGTGGCGCGCTCTGAGCAGAGCGCGCGCGCG 299
DB 307 TCATCGAGCGCTTACAGAGGTAGAGAGCGAGTGGCGCGCTCTGAGCAGAGCGCGCGCGCG 366
QY 300 TCCGCGAGCGCTTACAGAGGTAGAGAGCGAGTGGCGCGCTCTGAGCAGAGCGCGCGCGCG 359
DB 367 TCCGCGAGCGCTTACAGAGGTAGAGAGCGAGTGGCGCGCTCTGAGCAGAGCGCGCGCGCG 426
QY 360 ACCGCGAGCGCGCTTCTCTCTGCTGTGAGACGACGACGACGATGTAGTGTGTGCGCGCG 419
DB 427 ACCGCGAGCGCGCTTCTCTCTGCTGTGAGACGACGACGATGTAGTGTGTGCGCGCGCG 486
QY 420 CCATCGCGAGCATCTACGAGCGAGGAGAGAGAGAGAGCGCGCTTCTCTATATGGCTAGCG 479
DB 487 CCATCGCGAGCATCTACGAGCGAGGAGAGAGAGAGAGCGCGCTTCTCTATATGGCTAGCG 546
QY 480 CCGCGAGGAAACCTTGGCTTGTGAGCGAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCG 539
DB 547 CCGCGAGGAAACCTTGGCTTGTGAGCGAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
QY 540 GGGCGCGCGCTGAGCGCGCTGCGTACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
DB 607 GGGCGCGCGCTGAGCGCGCTGCGTACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 666

```

QY 600 TGTGGCTGGGAGGAGTGTGCCCCCTAGTCAGAGGCA 640  
DB 667 TGTGGCTGGGAGGAGTGTGCCCCCTAGTCAGAGGCA 707

## RESULT 4

US-09-471-275-1555

; Sequence 1555, Application US/09471275

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: Novel Contigs Obtained

; FILE REFERENCE: 782

; CURRENT FILING DATE: 1999-12-23

; EARLIER APPLICATION NUMBER: US 09/235,076

; EARLIER FILING DATE: 1999-01-20

; EARLIER APPLICATION NUMBER: US 09/234,611

; EARLIER FILING DATE: 1999-01-22

; EARLIER APPLICATION NUMBER: US 09/240,371

; EARLIER FILING DATE: 1999-01-29

; EARLIER APPLICATION NUMBER: US 09/277,227

; EARLIER FILING DATE: 1999-03-25

; EARLIER APPLICATION NUMBER: US 09/271,490

; EARLIER FILING DATE: 1999-03-18

; EARLIER APPLICATION NUMBER: US 09/293,972

; EARLIER FILING DATE: 1999-04-15

; EARLIER APPLICATION NUMBER: US 09/274,861

; EARLIER FILING DATE: 1999-03-23

; EARLIER APPLICATION NUMBER: US 60/125,453

; EARLIER FILING DATE: 1999-03-19

; EARLIER APPLICATION NUMBER: US 60/126,605

; EARLIER FILING DATE: 1999-03-26

; EARLIER APPLICATION NUMBER: US 09/306,350

; EARLIER FILING DATE: 1999-05-07

; EARLIER APPLICATION NUMBER: US 09/399,720

; EARLIER FILING DATE: 1999-09-21

; EARLIER APPLICATION NUMBER: US 09/404,284

; EARLIER FILING DATE: 1999-09-21

; EARLIER APPLICATION NUMBER:

; EARLIER FILING DATE: 1999-12-16

; NUMBER OF SEQ ID NOS: 10451

; SOFTWARE: pLCT\_genes Version 1.0

; SEQ ID NO 1555

; LENGTH: 977

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (155)...(517)

; OTHER INFORMATION: similar to g1455109 in the genepept database release 114,

; OTHER INFORMATION: Run with FASTX 3.3t00, default parameters

US-09-471-275-1555

Query Match 98.0%; Score 627.4; DB 21; Length 977;

Best Local Similarity 99.7%; Pred. No. 3.4e-107;

Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTCCCGGAGCGGAGCGGCTGCTCAGCGAGCGGCGGAGCGCTTACGCGGAGGCG 60  
DB 15 CTCCCGGAGCGGAGCGGCTGCTCAGCGAGCGGCGGAGCGCTTACGCGGAGGCG 74  
QY 61 GGAG-CGCCGAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 119  
DB 75 GGAGCGGCGGAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 134  
QY 120 CGCCGAGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 179  
DB 135 CGCCGAGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 194  
QY 180 CGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 239

DB 195 CGGAGCGGCTGAAGAGGAGTACAGCAGATCCGCGAGCAGCCAGCAACCAAAATCCCGGTGA 254  
QY 240 TCATGAGCGGCTACAGAGGAGTGAAGCAGCTGCCGCTCTGAGCAAGCAAGCTTTGG 299  
DB 255 TCATGAGCGGCTACAGAGGAGTGAAGCAGCTGCCGCTCTGAGCAAGCAAGCTTTGG 314  
QY 300 TCCCGAGCAGTCAACATGAGCGAGTGTGTCAGATCATCGGCGGCGCTGACGTGA 359  
DB 315 TCCCGAGCAGTCAACATGAGCGAGTGTGTCAGATCATCGGCGGCGCTGACGTGA 374  
QY 360 ACCGAGCAGGCGCTTTCTTCTGCTGTGTAACACAGCAGCATGTGAGTGTCCAGCC 419  
DB 375 ACCGAGCAGGCGCTTTCTTCTGCTGTGTAACACAGCAGCATGTGAGTGTCCAGCC 434  
QY 420 CCATGCGGAGCATACAGAGCAGGAAGACAGGAGCGGCTTCTATATGCTCTACG 479  
DB 435 CCATGCGGAGCATACAGAGCAGGAAGACAGGAGCGGCTTCTATATGCTCTACG 494  
QY 480 CTTCCAGGAAACCTTCGCTTCTGAGCCAGCAGTGAAGGAGGAGGAGGAGGAGGAG 539  
DB 495 CTTCCAGGAAACCTTCGCTTCTGAGCCAGCAGTGAAGGAGGAGGAGGAGGAGGAG 554  
QY 540 GGGCGGCGTCAAGCGGCTGCGCAGAGCCTTCTGCTTCTGAGTGAAGTGTCTACCG 599  
DB 555 GGGCGGCGTCAAGCGGCTGCGCAGAGCCTTCTGCTTCTGAGTGAAGTGTCTACCG 614  
QY 600 TGTGGCTGGGAGGAGTGTGCCCCCTAGTCAGAGGCA 640  
DB 615 TGTGGCTGGGAGGAGTGTGCCCCCTAGTCAGAGGCA 655

## RESULT 5

US-09-488-725B-8516

; Sequence 8516, Application US/09488725B

; GENERAL INFORMATION:

; APPLICANT: Yunnan T. Tang

; APPLICANT: John Tillinghast

; APPLICANT: Ankura Slinku

; APPLICANT: Chenghua Liu

; APPLICANT: Radoje T. Drmanac

; TITLE OF INVENTION: Novel Contigs Obtained

; FILE REFERENCE: 784

; CURRENT FILING DATE: 2000-01-21

; PRIOR APPLICATION NUMBER: US 09/004,182

; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: US 09/034,341

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 09/045,400

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: US 09/321,214

; PRIOR FILING DATE: 1999-05-26

; PRIOR APPLICATION NUMBER: US 09/131,598

; PRIOR FILING DATE: 1998-08-10

; PRIOR APPLICATION NUMBER: US 09/170,294

; PRIOR FILING DATE: 1998-10-13

; PRIOR APPLICATION NUMBER: US 09/179,473

; PRIOR FILING DATE: 1998-10-27

; PRIOR APPLICATION NUMBER: US 09/181,430

; PRIOR FILING DATE: 1998-10-28

; PRIOR APPLICATION NUMBER: US 09/235,076

; PRIOR FILING DATE: 1999-01-20

; PRIOR APPLICATION NUMBER: US 09/234,611

; Remaining Prior Application data removed - See File Wrapper or PALM.

; SOFTWARE: pLCT\_genes Version 1.01

; SEQ ID NO 8516

; LENGTH: 977

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:



NAME/KEY: misc.feature  
LOCATION: (155)...(517)  
OTHER INFORMATION: similar to q1455109 in the genepept database release 114,  
OTHER INFORMATION: Run with FASTX 3.3t00, default parameters  
US-09-488-725B-8516

Query Match 98.0%: Score 627.4; DB 21; Length 977;  
Best Local Similarity 99.7%: Pred. No. 3,4e-107;  
Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 CTCCCGCAGCCGCGACGCGCGTGTCTCAGCGCGAGCCCGCGAGCCCTTGAAGCGCGAGCGCG 60
DB 15 CTCCCGCAGCCGCGACGCGCGTGTCTCAGCGCGAGCCCGCGAGCCCTTGAAGCGCGAGCGCG 74
QY 61 GGAG-CGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 119
DB 75 GGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 134
QY 120 CGCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 145 CGCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 194
QY 180 CGGACCGCGTGAAGAGGATACAGAGATCCCGGACCGACCGCGCGCGCGCGCGCGCGCG 239
DB 195 CGGACCGCGTGAAGAGGATACAGAGATCCCGGACCGACCGCGCGCGCGCGCGCGCGCG 254
QY 240 TCATCGAGCGGTACAGAGGATGAGAGCGCTGCCGCTGCTGCGACAGCGCGCGCGCGCG 299
DB 255 TCATCGAGCGGTACAGAGGATGAGAGCGCTGCCGCTGCTGCGACAGCGCGCGCGCGCG 314
QY 300 TCCCGAGCGCATGTACACATGAGCGAGGTGTCAGATGATCCGGCGCGCGCGCGCGAGTGA 359
DB 315 TCCCGAGCGCATGTACACATGAGCGAGGTGTCAGATGATCCGGCGCGCGCGCGCGAGTGA 374
QY 360 ACCCGACGCGAGCGCTTCTCTGCTGCTGTAACACGACGACGATGGTGAAGTGTGTCACGC 419
DB 375 ACCCGACGCGAGCGCTTCTCTGCTGCTGTAACACGACGACGATGGTGAAGTGTGTCACGC 434
QY 420 CCATCGCGGACATGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 479
DB 435 CCATCGCGGACATGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 494
QY 480 CCTCCCGAGGAAACCTTGGGCTTCTGAGCGACGAGTAGGGGGGCTGGGCTGGAGTGGCG 539
DB 495 CCTCCCGAGGAAACCTTGGGCTTCTGAGCGACGAGTAGGGGGGCTGGGCTGGAGTGGCG 554
QY 540 GGGCGCGCGGTAGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 599
DB 555 GGGCGCGCGGTAGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 614
QY 600 TGGTGGGCTGGGCGAGCATGTGCCCGCTAGTCAAGAGGCA 640
DB 615 TGGTGGGCTGGGCGAGCATGTGCCCGCTAGTCAAGAGGCA 655
```

RESULT 6  
PCT-US01-18569-361  
Sequence 361, Application PC/TUS0118569

GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA133PCT  
CURRENT APPLICATION NUMBER: PCT/US01/18569  
CURRENT FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: 60/209,467  
PRIOR FILING DATE: 2000-06-07  
NUMBER OF SEQ ID NOS: 4360  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 361  
LENGTH: 1000  
TYPE: DNA  
ORGANISM: Homo sapiens  
PCT-US01-18569-361

Query Match 98.0%: Score 627.4; DB 1; Length 1000;  
Best Local Similarity 99.7%: Pred. No. 3,4e-107;  
Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 CTCCCGTAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 6 CTCCCGTAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 65
QY 61 GGAG-CGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 119
DB 66 GGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125
QY 120 CGCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 125 CGCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185
QY 180 CGGACCGCGTGAAGAGGATACAGAGATCCCGGACCGACCGCGCGCGCGCGCGCGCGCG 239
DB 186 CGGACCGCGTGAAGAGGATACAGAGATCCCGGACCGACCGCGCGCGCGCGCGCGCGCG 245
QY 240 TCATCGAGCGGTACAGAGGATGAGAGCGCTGCCGCTGCTGCGACAGCGCGCGCGCGCG 299
DB 246 TCATCGAGCGGTACAGAGGATGAGAGCGCTGCCGCTGCTGCGACAGCGCGCGCGCGCG 305
QY 300 TCCCGAGCGCATGTACACATGAGCGAGGTGTCAGATGATCCGGCGCGCGCGCGCGAGTGA 359
DB 306 TCCCGAGCGCATGTACACATGAGCGAGGTGTCAGATGATCCGGCGCGCGCGCGCGAGTGA 365
QY 360 ACCCGACGCGAGCGCTTCTCTGCTGCTGTAACACGACGACGATGGTGAAGTGTGTCACGC 419
DB 366 ACCCGACGCGAGCGCTTCTCTGCTGCTGTAACACGACGACGATGGTGAAGTGTGTCACGC 425
QY 420 CCATCGCGGACATGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 479
DB 426 CCATCGCGGACATGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 485
QY 480 CCTCCCGAGGAAACCTTGGGCTTCTGAGCGACGAGTAGGGGGGCTGGGCTGGAGTGGCG 539
DB 486 CCTCCCGAGGAAACCTTGGGCTTCTGAGCGACGAGTAGGGGGGCTGGGCTGGAGTGGCG 545
QY 540 GGGCGCGCGGTAGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 599
DB 546 GGGCGCGCGGTAGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 605
QY 600 TGGTGGGCTGGGCGAGCATGTGCCCGCTAGTCAAGAGGCA 640
DB 606 TGGTGGGCTGGGCGAGCATGTGCCCGCTAGTCAAGAGGCA 646
```

RESULT 7  
PCT-US01-18569-361  
Sequence 361, Application PC/TUS0118569

GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA133PCT  
CURRENT APPLICATION NUMBER: PCT/US01/18569  
CURRENT FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: 60/209,467  
PRIOR FILING DATE: 2000-06-07  
NUMBER OF SEQ ID NOS: 4360  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 361  
LENGTH: 1000  
TYPE: DNA  
ORGANISM: Homo sapiens  
PCT-US01-18569-361

Query Match 98.0%: Score 627.4; DB 2; Length 1000;  
Best Local Similarity 99.7%: Pred. No. 3,4e-107;  
Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

0Y 1 CTCCGCGAGCGAGCGCGCTGCTCAGCGGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 60  
DB 6 CTCCGCGAGCGAGCGCGCGCTGCTCAGCGGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 65  
0Y 61 GGAG-CCCCGGAGCCCCCAAAACCGAGACATCCCGCGCCCGGAGAGCCCGGCTGCG 119  
DB 66 GGAGCGCCCGGAGCGCCCGGAGAGCGAGACATCCCGCGCCCGGAGAGCCCGGCTGCG 125  
0Y 120 CGCCGAGCGCGCGCGCGATGCGCTCAGACCGCGCTTTTGAAGAGCGCGGAGCTTGC 179  
DB 126 CGCCGAGCGCGCGCGCGCGATGCGCTCAGACCGCGCTTTTGAAGAGCGCGGAGCTTGC 185  
0Y 180 CGGACCGCTGTAAAGAGGTACAGCAGATCCCGGAGCAGCCCGGAGAAATCCCGTGA 239  
DB 186 CGGACCGCTGTAAAGAGGTACAGCAGATCCCGGAGCAGCCCGGAGAAATCCCGTGA 245  
0Y 240 TCATGAGCGCTACAGAGGTGAGAGCAGTCCCGCTCTGAGACAGCAAGTTTGG 299  
DB 246 TCATGAGCGCTACAGAGGTGAGAGCAGTCCCGCTCTGAGACAGCAAGTTTGG 305  
0Y 300 TCCCGAGCATGTCAACATGAGCGAGTGTCAAGATCATCCGCGCGCGCTGACGTGA 359  
DB 306 TCCCGAGCATGTCAACATGAGCGAGTGTCAAGATCATCCGCGCGCGCTGACGTGA 365  
0Y 360 ACCCGAGCAGCGCTTTCTTCTGCTGTGAACAGCAGCAGATGCTGATGCTCCAGCG 419  
DB 366 ACCCGAGCAGCGCTTTCTTCTGCTGTGAACAGCAGCAGATGCTGATGCTCCAGCG 425  
0Y 420 CCATGCGGACATCTACAGAGCAGAGAAAGAGAGCGGCTCTCTATATATGCTGACG 479  
DB 426 CCATGCGGACATCTACAGAGCAGAGAAAGAGAGCGGCTCTCTATATATGCTGACG 485  
0Y 480 CCTCCAGAGAAACCTTTCGCTTCTGAGCAGCAGATGAGGGGCTTCGCTGAGTCCGG 539  
DB 486 CCTCCAGAGAAACCTTTCGCTTCTGAGCAGCAGATGAGGGGCTTCGCTGAGTCCGG 545  
0Y 540 GGGCGCGCGTACAGCGCTGCGCCAGAGAGCTTCTGTTCTTGAACCTGAGCTGCTTACCG 599  
DB 546 GGGCGCGCGTACAGCGCTGCGCCAGAGAGCTTCTGTTCTTGAACCTGAGCTGCTTACCG 605  
0Y 600 TGTGCGCTGGGAGCAGCATGTGCCCTTACTAGCAGAGGCA 640  
DB 606 TGTGCGCTGGGAGCAGCATGTGCCCTTACTAGCAGAGGCA 646

RESULT 8  
US-10-264-049-361  
; Sequence 361, Application US/10264049  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA13PI  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 361  
; LENGTH: 1000  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-264-049-361

Query Match 98.0%; Score 627.4; DB 49; Length 1000;  
Best Local Similarity 99.7%; Pred. No. 3.4e-107;  
Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

0Y 1 CTCCGCGAGCGAGCGCGCTGCTCAGCGGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 60  
DB 6 CTCCGCGAGCGAGCGCGCGCTGCTCAGCGGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 65

0Y 61 GGAG-CCCCGGAGCCCCCAAAACCGAGACATCCCGCGCCCGGAGAGCCCGGCTGCG 119  
DB 66 GGAGCGCCCGGAGCGCCCGGAGAGCGAGACATCCCGCGCCCGGAGAGCCCGGCTGCG 125  
0Y 120 CGCCGAGCGCGCGCGCGATGCGCTCAGACCGCGCTTTTGAAGAGCGCGGAGAGCTTGC 179  
DB 126 CGCCGAGCGCGCGCGCGATGCGCTCAGACCGCGCTTTTGAAGAGCGCGGAGAGCTTGC 185  
0Y 180 CGGACCGCTGTAAAGAGGTACAGCAGATCCCGGAGCAGCCCGGAGAAATCCCGTGA 239  
DB 186 CGGACCGCTGTAAAGAGGTACAGCAGATCCCGGAGCAGCCCGGAGAAATCCCGTGA 245  
0Y 240 TCATGAGCGCTACAGAGGTGAGAGCAGTCCCGCTCTGAGACAGCAAGTTTGG 299  
DB 246 TCATGAGCGCTACAGAGGTGAGAGCAGTCCCGCTCTGAGACAGCAAGTTTGG 305  
0Y 300 TCCCGAGCATGTCAACATGAGCGAGTGTCAAGATCATCCGCGCGCGCTGACGTGA 359  
DB 306 TCCCGAGCATGTCAACATGAGCGAGTGTCAAGATCATCCGCGCGCGCTGACGTGA 365  
0Y 360 ACCCGAGCAGCGCTTTCTTCTGCTGTGAACAGCAGCAGATGCTGATGCTCCAGCG 419  
DB 366 ACCCGAGCAGCGCTTTCTTCTGCTGTGAACAGCAGCAGATGCTGATGCTCCAGCG 425  
0Y 420 CCATGCGGACATCTACAGAGCAGAGAAAGAGAGCGGCTCTCTATATATGCTGACG 479  
DB 426 CCATGCGGACATCTACAGAGCAGAGAAAGAGAGCGGCTCTCTATATATGCTGACG 485  
0Y 480 CCTCCAGAGAAACCTTTCGCTTCTGAGCAGCAGATGAGGGGCTTCGCTGAGTCCGG 539  
DB 486 CCTCCAGAGAAACCTTTCGCTTCTGAGCAGCAGATGAGGGGCTTCGCTGAGTCCGG 545  
0Y 540 GGGCGCGCGTACAGCGCTGCGCCAGAGAGCTTCTGTTCTTGAACCTGAGCTGCTTACCG 599  
DB 546 GGGCGCGCGTACAGCGCTGCGCCAGAGAGCTTCTGTTCTTGAACCTGAGCTGCTTACCG 605  
0Y 600 TGTGCGCTGGGAGCAGCATGTGCCCTTACTAGCAGAGGCA 640  
DB 606 TGTGCGCTGGGAGCAGCATGTGCCCTTACTAGCAGAGGCA 646

RESULT 9  
US-60-234-690-7194  
; Sequence 7194, Application US/60234690  
; GENERAL INFORMATION:  
; APPLICANT: Morris; Macdonald  
; APPLICANT: Diep, Dinh  
; TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING  
; FILE REFERENCE: POLYNUCLEOTIDE SEQUENCE DATABASES, AND SINGLE NUCLEOTIDE  
; CURRENT APPLICATION NUMBER: US/60/234,690  
; CURRENT FILING DATE: 2000-09-21  
; NUMBER OF SEQ ID NOS: 12060  
; SOFTWARE: PERL Program  
; SEQ ID NO 7194  
; LENGTH: 1009  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 247776.13  
US-60-234-690-7194

Query Match 98.0%; Score 627.4; DB 77; Length 1009;  
Best Local Similarity 99.7%; Pred. No. 3.4e-107;  
Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

0Y 1 CTCCGCGAGCGAGCGCGCGCTGCTCAGCGGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 60  
DB 57 CTCCGCGAGCGAGCGCGCGCTGCTCAGCGGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 116

QY	61	GAG-CCCCGAGCCCCCAACCGCAGACATCCGCGCGCCGAGAGCCGCTTGG	119
Lb	117		
QY	120	CGCCGAGCCGGGCCCGCGCGATGCCCTAGACCGGCGCTTTCAAGCAGCGCGAGCTTG	179
Db	177	CGCCGAGCCGGGCCCGCGCGATGCCCTAGACCGGCGCTTTCAAGCAGCGCGAGCTTG	236
QY	180	CCGACCGCTGTAAAGAGGTACAGCAGATCCGCGACAGCAGCCCAACAAATCCCGTGA	239
Db	237	CCGACCGCTGTAAAGAGGTACAGCAGATCCGCGACAGCAGCCCAACAAATCCCGTGA	296
QY	240	TCATGAGCGCTACAAAGGTGAGAGCAGCTGCCGCTCTGGACAAAGCAAGTTTGG	299
Lb	297	TCATGAGCGCTACAAAGGTGAGAGCAGCTGCCGCTCTGGACAAAGCAAGTTTGG	356
QY	300	TCGCGGACCATGTCAACATGAGCGAGTTGGTCAAGATCATCCGCGCGCCGCTGCAGCTGA	359
Db	357	TCGCGGACCATGTCAACATGAGCGAGTTGGTCAAGATCATCCGCGCGCCGCTGCAGCTGA	416
QY	417	ACCCGAGCGAGGCGCTTCTCTGCTGGTGAACACACAGCATGGTGAATGGTCCAGCG	419
Lb	417		
QY	420	CCATGCGCGAGCATTCAGCAGCAGAGAAAGAGCAGAGCGCTTCCATATATGGTTAG	479
Lb	477	CCATGCGCGAGCATTCAGCAGCAGAGAAAGAGCAGAGCGCTTCCATATATGGTTAG	536
QY	480	CCTCCGAGAAACCTTGGCGTTCTGAGCCAGCATAGAGGGGCGCTCGGCGTGGAGTGGG	539
Db	537	CCTCCGAGAAACCTTGGCGTTCTGAGCCAGCATAGAGGGGCGCTCGGCGTGGAGTGGG	596
QY	540	GCGCGCCCGGTAGGCGCTGCCAGAGCTTGTGCTCTGAAGTAGAGTGCCTTAGCG	599
Db	597	GCGCGCCCGGTAGGCGCTGCCAGAGAGCTGTGGTCTCTGAAGTAGAGTGCCTTAGCG	656
QY	600	TGTGTGGCTGGGCGAGCATGTGCCCGCTAGTCAGAGGGCA	640
Db	657	TGTGTGGCTGGGCGAGCATGTGCCCGCTAGTCAGAGGGCA	697

RESULT 10

US-60-323-843-7194

Sequence 7194, Application US/60323843

GENERAL INFORMATION:

APPLICANT: Morris, Macdonald

APPLICANT: Lal, Preeti

APPLICANT: Deep, Dinu

TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING

TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCE DATABASES, AND SINGLE NUCLEOTIDE

FILE REFERENCE: GX-0021-1 P

CURRENT APPLICATION NUMBER: US/60/323,843

CURRENT FILING DATE: 2001-09-21

NUMBER OF SEQ ID NOS: 12060

SOFTWARE: PERL Program

SEQ ID NO 7194

LENGTH: 1009

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID NO: 247776.13

US-60-323-843-7194

Query Match	98.0%	Score 627.4	DB 87	Length 1009
Best Local Similarity	Pred. 90.7%	Pred. 3.4e-107		
Matches 639	Conservative 0	Mismatches 1	Indels 1	Gaps 1

  

QY	1	CTCCCCAGCCGCGCCGCTGCTCAAGCGTGAAGCCCGCGAGACCTTTAGTGGCCAGGCTGC	60
Db	57	CTCCCGGAGCCGCGACGCGCGTGTCTACAGCGGAGGCCGCGAGCCCTTATAGCGCCAGGGCC	116

QY	61	GGAG-CCGCGGAGCCCTTAAACCGAGACATCCCGGCGCCAGAGCCCGGCTTCC	119
		TT	
Db	117	GGAGCCCCCGGAGCCCGCCAAACCGAGACATCCCGGCGCCAGAGCCCGGCTTCC	176
QY	120	CGCCAGCGCGGCGCCCGCGGATGCGCTTCGAGCGGCGCTTTCGAAGCGAGCGGCTTCG	179
		CC	
Db	177	CGCCAGCGGCGCGCCCGCGGATGCGCTTTCGAGCGGCGCTTTCGAAGCGAGCGGCTTCG	236
QY	180	CGAGCCCGCTGTAAAGAGGATACAGCAATATCTGCGACCGACCCCGCAATATCTGCTCA	239
Db	237	CGAGCCCGCTGTAAAGAGGATACAGCAATATCTGCGACCGACCCCGCAATATCTGCTCA	296
QY	240	TCATTCAGCGGCTACCAAGGCTGAGAGAGTACGTGCGCGCTCTCGAGCAAGACATAATTTT	299
Db	297	TCATTCAGCGGCTACCAAGGCTGAGAGAGTACGTGCGCGCTCTCGAGCAAGACATAATTTT	356
QY	300	TCGCGGACCATTTCAACATGAGCGAGTTGCTTAAGATCATCGCGCGCGCGCTTCAATTA	359
Db	357	TCGCGGACCATTTCAACATGAGCGAGTTGCTTAAGATCATCGCGCGCGCGCTTCAATTA	416
QY	360	ACCCGAGCGAGCGCTTCTTCTGCTGGTGAACCCAGACAGCATGCTGAGTGTATCTATCTAT	419
		TT	
Db	417	ACCCGAGCGAGCGCTTCTTCTGCTGGTGAACCCAGACAGCATGCTGAGTGTATCTATCTAT	476
QY	420	CCATTGCGGACATCTTACGAGCGAGAGAAAGAGAGACGCGCTCTCTATATATATATATAT	479
Db	477	CCATGCGCGGACATCTTACGAGCGAGAGAAAGAGAGACGCGCTCTCTATATATATATATAT	536
QY	480	CTCCGAGGAACCTTCGCGCTCTGAGCGGACAGTAGTAGGGGGGGCTCGGGCTGGTAAATGAG	539
		TT	
Db	537	CTCCGAGGAACCTTCGCGCTCTGAGCGGACAGTAGTAGGGGGGGCTCGGGCTGGTAAATGAG	596
QY	540	GGGCGCCCGGTGAGGCCCTGCCAGAGAGCTCTGCTTCCTGAACTGAGCTGCTCTATATAT	599
		TT	
Db	597	GGGCGCCCGGTGAGGCCCTGCCAGAGAGCTCTGCTTCCTGAACTGAGCTGCTCTATATAT	656
QY	600	TGCTGGGCGTGGCGAGGAGATGTGGCGGCTTAGTCAGAGGCTA	640
		TT	
Db	657	TGCTGGGCGTGGCGAGGAGATGTGGCGGCTTAGTCAGAGGCTA	697

```

? RESULT 11
? US-10-405-027-30
? Sequence 30: Application US/10405027
? GENERAL INFORMATION:
? APPLICANT: Rosen et al.
? TITLE OF INVENTION: Human Secreted Proteins
? FILE REFERENCE: PS806P1
? CURRENT APPLICATION NUMBER: US/10/405,027
? CURRENT FILING DATE: 2003-04-07
? PRIOR APPLICATION NUMBER: 60/359,608
? PRIOR FILING DATE: 2002-04-04
? PRIOR APPLICATION NUMBER: 60/376,175
? PRIOR FILING DATE: 2002-04-30
? NUMBER OF SEQ ID NOS: 5810
? SOFTWARE: PatentIn Ver. 2.10
? SFO ID NO 30
? LENGTH: 1019
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-10-405-027-30

```

Query Match	Similarity	98.0%	Score	627.4	DB	52	Length	1019	
Best Local	Similarity	99.7%	Pred. No.	3,4e-107					
Matches	639	Conservative	0	Mismatches	1	Indels	1	Gaps	1

  

QY	1	CTCCGCTGAGCCGACAGTCTTCTGTATCTTCTGATCCCGCGAGTCTTATGAGTGTGATCTCTTG	60
Db	25	CTCCGCGAGCCCGACGCTGCTGCTCTCATATGATGAGCCCGCGAAGGCTTGAATGCGATATCTTGT	84
QY	61	GGAGCTCTGGGAGTCTCTAAATCGATATATATCCCGCTGCTGCTGATGAGGCTGCTCTCTCTCT	119

Db 85 GGAGCCCCGGAGCCCCCAAAACCGAGACATCCCGCCGCCAGAGCCCCGGCCCTCG 144  
QY 120 CGCCAGCCGGGCCCCCGCGGATGCCCTGAGACCGGCTTTTCAAGAGCGCGGAGCTTCG 179  
Db 145 CGCCAGCCGGGCCCCCGCGGATGCCCTGAGACCGGCTTTTCAAGAGCGCGGAGCTTCG 204  
QY 180 CGAGCCGCTGTAAAGAGGTAGACAGATCCGGGACAGACCCCAAGCAAAATCCGGTGA 239  
Db 205 CGAGCCGCTGTAAAGAGGTAGACAGATCCGGGACAGACCCCAAGCAAAATCCGGTGA 264  
QY 240 TCATGAGACGCTACACAGGAGTGAAGAGAGCTGCCGCTCTGAGACAAGACCAAGTTTGG 299  
Db 265 TCATGAGACGCTACACAGGAGTGAAGAGAGCTGCCGCTCTGAGACAAGACCAAGTTTGG 324  
QY 300 TCCCGGACCATGTCAACATGAGCGAGTGTCAAGATATCCGGGCGCCCTGACAGCTGA 359  
Db 325 TCCCGGACCATGTCAACATGAGCGAGTGTCAAGATATCCGGGCGCCCTGACAGCTGA 384  
QY 360 ACCCCAGCAGGCTTTCTTCTGCTGTGAACCAAGACAGCATGGTGTGTCCAGCG 419  
Db 385 ACCCCAGCAGGCTTTCTTCTGCTGTGAACCAAGACAGCATGGTGTGTCCAGCG 444  
QY 420 CCATCGCGGACATCTACAGAGCAGAGAGAGAGAGAGCGCTCTCTATATGCTACG 479  
Db 445 CCATCGCGGACATCTACAGAGCAGAGAGAGAGAGAGCGCTCTCTATATGCTACG 504  
QY 480 CCTCCAGGAACCTTGGCTTCTGAGCAGCAGTGAAGGAGGCTGCGCTGGAGTCCGG 539  
Db 505 CCTCCAGGAACCTTGGCTTCTGAGCAGCAGTGAAGGAGGCTGCGCTGGAGTCCGG 564  
QY 540 GGGCCCGGCTGAGGCTTCCGCGAGAGAGCTTGTGTTCTTAAGTGAAGCTGCTTACG 599  
Db 565 GGGCCCGGCTGAGGCTTCCGCGAGAGAGCTTGTGTTCTTAAGTGAAGCTGCTTACG 624  
QY 600 TGTGTGGCTGGGAGGATGTGCCCCCTAGTCAGAGGGCA 640  
Db 625 TGTGTGGCTGGGAGGATGTGCCCCCTAGTCAGAGGGCA 665

## RESULT 12

US-09-428-151A-9451  
: Sequence 9451, Application US/09428151A  
: GENERAL INFORMATION:  
: APPLICANT: Glucksmann, M. Alexandra  
: TITLE OF INVENTION: Nucleic Acid Molecules Derived from a  
: FILE REFERENCE: 1600.1003001  
: CURRENT APPLICATION NUMBER: US/09/428,151A  
: PRIOR FILING DATE: 1999-10-27  
: PRIOR APPLICATION NUMBER: 60/106,057  
: NUMBER OF SEQ ID NOS: 11036  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 9451  
: LENGTH: 1356  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
US-09-428-151A-9451

Query Match 98.0%; Score 627.4; DB 21; Length 1356;  
Best Local Similarity 99.7%; Pred. No. 3.4e-107;

Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTCCCGAGCGCGAGCGCGCTGCTCAAGCGGAGCGCCCGAGCCCTTGAAGCGAGGGCG 60  
Db 391 CTCCCGAGCGCGAGCGCGCGCTGCTCAAGCGGAGCGCCCGAGCCCTTGAAGCGAGGGCG 450  
QY 61 GGAG-CCCGGAGCGCCCAAAACCGAGACATCCCGCGCCCAAGAGCCCGGCTTCG 119  
Db 451 GGAGCGCGCGGAGCGCCCAAAACCGAGACATCCCGCGCCCAAGAGCGCCCGGCTTCG 510  
QY 120 CGCCAGCGCGGCGCGGCGATGCCCTTCAAGCGGCTTTTAACGAGCGGCGAGCTTCG 179  
Db 120 CGCCAGCGCGGCGCGGCGATGCCCTTCAAGCGGCTTTTAACGAGCGGCGAGCTTCG 179

Db 511 CGCCAGCGGCGCGCGCGGATGCTCAGACCGGCTTTTCAAGCAGCGCGGAGCTTCG 570  
QY 160 CGAGCGCTGTAAAGAGGTAGACAGATCCGGGACAGACCCCAAGCAAAATCCGGTGA 239  
Db 571 CGAGCGCTGTAAAGAGGTAGACAGATCCGGGACAGACCCCAAGCAAAATCCGGTGA 630  
QY 240 TCATGAGACGCTACACAGGAGTGAAGAGAGCTGCCGCTCTGAGACAAGACCAAGTTTGG 299  
Db 621 TCATGAGACGCTACACAGGAGTGAAGAGAGCTGCCGCTCTGAGACAAGACCAAGTTTGG 690  
QY 300 TCCCGGACCATGTCAACATGAGCGAGTGTCAAGATATCCGGGCGCCCTGACAGCTGA 359  
Db 691 TCCCGGACCATGTCAACATGAGCGAGTGTCAAGATATCCGGGCGCCCTGACAGCTGA 750  
QY 360 ACCCCAGCAGGCTTTCTTCTGCTGTGAACCAAGACAGCATGGTGTGTCCAGCG 419  
Db 751 ACCCCAGCAGGCTTTCTTCTGCTGTGAACCAAGACAGCATGGTGTGTCCAGCG 810  
QY 420 CCATCGCGGACATCTACAGAGCAGAGAGAGAGAGAGCGCTCTCTATATGCTACG 479  
Db 811 CCATCGCGGACATCTACAGAGCAGAGAGAGAGAGAGCGCTCTCTATATGCTACG 870  
QY 480 CCTCCAGGAACCTTGGCTTCTGAGCAGCAGTGAAGGAGGCTGCGCTGGAGTCCGG 539  
Db 871 CCTCCAGGAACCTTGGCTTCTGAGCAGCAGTGAAGGAGGCTGCGCTGGAGTCCGG 930  
QY 540 GGGCCCGGCTGAGGCTTCCGCGAGAGAGCTTGTGTTCTTAAGTGAAGCTGCTTACG 599  
Db 931 GGGCCCGGCTGAGGCTTCCGCGAGAGAGCTTGTGTTCTTAAGTGAAGCTGCTTACG 990  
QY 600 TGTGTGGCTGGGAGGATGTGCCCCCTAGTCAGAGGGCA 640  
Db 991 TGTGTGGCTGGGAGGATGTGCCCCCTAGTCAGAGGGCA 1021

## RESULT 13

PCT-US98-15949-94  
: Sequence 94, Application PC/TUS9815949  
: GENERAL INFORMATION:  
: APPLICANT: Human Genome Sciences, Inc., et al.  
: TITLE OF INVENTION: 83 Human Secreted Proteins  
: FILE REFERENCE: P2012PCT  
: CURRENT APPLICATION NUMBER: PCT/US98/15949  
: PRIOR FILING DATE: 1998-07-29  
: NUMBER OF SEQ ID NOS: 219  
: SOFTWARE: PatentIn Ver. 2.0  
: SEQ ID NO 94  
: LENGTH: 951  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: SITE  
: LOCATION: (826)  
: OTHER INFORMATION: n equals a,t,g, or c  
PCT-US98-15949-94

Query Match 98.0%; Score 627; DB 1; Length 951;  
Best Local Similarity 99.5%; Pred. No. 4.1e-107;

Matches 638; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTCCCGAGCGCGAGCGCGCTGCTCAAGCGGAGCGCCCGAGCCCTTGAAGCGAGGGCG 60  
Db 6 CTCCCGAGCGCGAGCGCGCGCTGCTCAAGCGGAGCGCCCGAGCCCTTGAAGCGAGGGCG 65  
QY 61 GGAG-CCCGGAGCGCCCAAAACCGAGACATCCCGCGCCCAAGAGCCCGGCTTCG 119  
Db 66 GGAGCGCGCGGAGCGCCCAAAACCGAGACATCCCGCGCCCAAGAGCCCGGCTTCG 125  
QY 120 CGCCAGCGGCGCGGCGGCGATGCCCTTCAAGCGGCGCTTTTAACGAGCGGCGAGCTTCG 179  
Db 126 CGCCAGCGGCGCGGCGGCGATGCCCTTCAAGCGGCGCTTTTAACGAGCGGCGAGCTTCG 185  
QY 180 CGAGCGCTGTAAAGAGGTAGACAGATCCCGAGACAGACCCCAAGCAAAATCCGGTGA 239

```
|||||
186 CCGACCGCTGTAAAGAGGTACAGCAGATCCGGAGCAGACCCAGCAAAATCCGGTGA 245
240 TCATCGACCGCTACAAAGGTGAGAGCAGAGCTCCCGCTCTGGACAGATCAAGTTTGG 299
246 TCATCGACCGCTACAAAGGTGAGAGCAGAGCTCCCGCTCTGGACAGATCAAGTTTGG 305
400 TCCCGGACCATGTCAACATGAGACGAGTTGTCAAGATCATCGCGCGCGCTGCAGCTGA 359
406 TCCCGGACCATGTCAACATGAGACGAGTTGTCAAGATCATCGCGCGCGCTGCAGCTGA 365
460 ACCCGACGAGCGCTTCTTCTGCTGTGTGTAACCCAGCAGACAGATGCTGATGTCACAC 419
466 ACCCGACGAGCGCTTCTTCTGCTGTGTGTAACCCAGCAGACAGATGCTGATGTCACAC 425
420 CCATCGCGGACATCTACAGACAGAGAGAAAGACAGAGCGCTTCTCTATATAGTCTACG 479
426 CCATCGCGGACATCTACAGACAGAGAGAAAGACAGAGCGCTTCTCTATATAGTCTACG 485
480 CCTCCAGAGAAACCTTCCGCTTCTGAGCCAGCAGATAGGGGGCTCGGCTGGAGTCCGG 539
486 CCTCCAGAGAAACCTTCCGCTTCTGAGCCAGCAGATAGGGGGCTCGGCTGGAGTCCGG 545
540 GGGCGCGGCTAGGCGCTGCGCCAGAGAGCTTCTGATTCCTGAATGAGTGCCTTACCG 599
546 GGGCGCGGCTAGGCGCTGCGCCAGAGAGCTTCTGATTCCTGAATGAGTGCCTTACCG 605
600 TGGTGGGCTGGGAGGCAATGCGCCGCTAGTCACAGGGCA 640
606 TGGTGGGCTGGGAGGCAATGCGCCGCTAGTCACAGGGCA 646
```

## RESULT 14

PCT-US98-1549-94

```
: Sequence 94, Application PC/TUS9815949
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc., et al.
: TITLE OF INVENTION: 83 Human Secreted Proteins
: FILE REFERENCE: P2012PCT
: CURRENT APPLICATION NUMBER: PCT/US98/15949
: CURRENT FILING DATE: 1998-07-29
: NUMBER OF SEQ ID NOS: 219
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 94
: LENGTH: 951
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (826)
: OTHER INFORMATION: n equals a,t,g, or c
PCT-US98-1549-94
```

Query Match 98.0%; Score 627; DB 2; Length 951;  
Best Local Similarity 99.5%; Pred. No. 4.1e-107;  
Matches 638; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

```
1 CTCGGGAGCGCGGAGCGCGCGCTGCTCAGCGGAGCGCGCGGAGCGCTTACGCGGAGCGG 60
|||||
6 CTCGGGAGCGCGGAGCGCGCGCTGCTCAGCGGAGCGCGCGGAGCGCTTACGCGGAGCGG 65
61 GGAG-CCCGGAGCGCGCGGAGCGCGGAGCGGAGCGCGCGGAGCGCGCGGAGCGCGCTG 119
|||||
66 GGAGCGCGCGGAGCGCGCGGAGCGCGGAGCGGAGCGCGCGGAGCGCGCGGAGCGCGCTG 125
120 GCGCGAGCGCGCGCGCGGAGCGCGGAGCGCGCGGAGCGCGCGGAGCGCGCGGAGCGCG 179
|||||
126 GCGCGAGCGCGCGCGCGGAGCGCGGAGCGCGCGGAGCGCGCGGAGCGCGCGGAGCGCG 185
180 CCGAGCGGCTGTAAAGAGGTACAGCAGATCCGGAGCAGACCCAGCAAAATCCGGTGA 239
186 CCGAGCGGCTGTAAAGAGGTACAGCAGATCCGGAGCAGACCCAGCAAAATCCGGTGA 245
```

```
240 TCATCGACCGCTACAAAGGTGAGAGCAGAGCTCCCGCTCTGGACAGATCAAGTTTGG 299
246 TCATCGACCGCTACAAAGGTGAGAGCAGAGCTCCCGCTCTGGACAGATCAAGTTTGG 305
300 TCCCGGACCATGTCAACATGAGACGAGTTGTCAAGATCATCGCGCGCGCTGCAGCTGA 359
306 TCCCGGACCATGTCAACATGAGACGAGTTGTCAAGATCATCGCGCGCGCTGCAGCTGA 365
360 ACCCGACGAGCGCTTCTTCTGCTGTGTGTAACCCAGCAGACAGATGCTGATGTCACAC 419
366 ACCCGACGAGCGCTTCTTCTGCTGTGTGTAACCCAGCAGACAGATGCTGATGTCACAC 425
420 CCATCGCGGACATCTACAGACAGAGAGAAAGACAGAGCGCTTCTCTATATAGTCTACG 479
426 CCATCGCGGACATCTACAGACAGAGAGAAAGACAGAGCGCTTCTCTATATAGTCTACG 485
480 CCTCCAGAGAAACCTTCCGCTTCTGAGCCAGCAGATAGGGGGCTCGGCTGGAGTCCGG 539
486 CCTCCAGAGAAACCTTCCGCTTCTGAGCCAGCAGATAGGGGGCTCGGCTGGAGTCCGG 545
540 GGGCGCGGCTAGGCGCTGCGCCAGAGAGCTTCTGATTCCTGAATGAGTGCCTTACCG 599
546 GGGCGCGGCTAGGCGCTGCGCCAGAGAGCTTCTGATTCCTGAATGAGTGCCTTACCG 605
600 TGGTGGGCTGGGAGGCAATGCGCCGCTAGTCACAGGGCA 640
606 TGGTGGGCTGGGAGGCAATGCGCCGCTAGTCACAGGGCA 646
```

## RESULT 15

US-60-172-373-10400

```
: Sequence 10400, Application US/6017273
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lal, Preeti
: APPLICANT: Deep, Dink
: TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
: FILE REFERENCE: GX-000619
: CURRENT APPLICATION NUMBER: US/60/172,373
: CURRENT FILING DATE: 1999-12-16
: NUMBER OF SEQ ID NOS: 25,772
: SOFTWARE: PERL Program
: SEQ ID NO: 10400
: LENGTH: 1017
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incycle ID No: 247776.3
: NAME/KEY: unsure
: LOCATION: 14, 25, 46
: OTHER INFORMATION: a, t, c, g, or other
US-60-172-373-10400
```

Query Match 94.6%; Score 605.4; DB 71; Length 1017;  
Best Local Similarity 99.4%; Pred. No. 4.3e-103;  
Matches 639; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

```
1 CTCGGGAGCGCGGAGCGCGCGCTGCTCAGCGGAGCGCGCGGAGCGCTTACGCGGAGCGG 60
|||||
71 CTCGGGAGCGCGGAGCGCGCGCTGCTCAGCGGAGCGCGCGGAGCGCTTACGCGGAGCGG 140
61 GGAG-CCCGGAGCGCGCGGAGCGCGGAGCGGAGCGCGCGGAGCGCGCGGAGCGCGCTG 117
|||||
131 GGAGCGCGCGGAGCGCGCGGAGCGCGGAGCGGAGCGCGCGGAGCGCGCGGAGCGCGCTG 190
118 GCGCGGAGCGCGCGCGCGGAGCGCGGAGCGCGGAGCGCGCGGAGCGCGCGGAGCGCG 177
|||||
191 GCGCGGAGCGCGCGCGCGGAGCGCGGAGCGCGGAGCGCGCGGAGCGCGCGGAGCGCG 250
178 GCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAG 247
```

```

Db      251 |||||CGCCACCGCTGTAGAGGTACAGCAGATCCGACACGACGCCAGCAAAATCCCGGT 310
QY      238 GATCATCGAGCGCTCAAGGGTGAAGAGAGCTGCCGCTCTGGACAAAGCAACTTTT 297
Db      311 GATCATCGAGCGCTCAAGGGTGAAGAGAGCTGCCGCTCTGGACAAAGCAAGTTT 370
QY      298 GGTCCCGGACCATGTCAACATGAGCGAGTTGGTCAAGATCATCCGGCCGCGCTGCAGCT 357
Db      371 GGTCCCGGACCATGTCAACATGAGCGAGTTGGTCAAGATCATCCGGCCGCGCTGCAGCT 430
QY      358 GAACCCCAAGCAGCGCTTCTCTGCTGCTGTAACCAAGCAGCAGCATGTGAGTGTCCAC 417
Db      431 GAACCCCAAGCAGCGCTTCTCTGCTGCTGTAACCAAGCAGCAGCATGTGAGTGTCCAC 490
QY      418 GCCCATCGGGACATCTAGAGAGAGAAAGAGAGAGAGCGGCTCTCTATATGTTCTA 477
Db      491 GCCCATCGGGACATCTAGAGAGAGAAAGAGAGAGAGCGGCTCTCTATATGTTCTA 550
QY      478 CGCCTCCAGGAACCTTCGGCTTCTGAGCCAGCAGTAGAGGGGCTCGGCTGGAGTGG 537
Db      551 CGCCTCCAGGAACCTTCGGCTTCTGAGCCAGCAGTAGAGGGGCTCGGCTGGAGTGG 610
QY      538 GGGGGCCCCGTCAGGGCCCTGCCAGAGAGCTTCTGTTCTTGAAGTGAAGTGCCTTAC 597
Db      611 GGGGGCCCCGTCAGGGCCCTGCCAGAGAGCTTCTGTTCTTGAAGTGAAGTGCCTTAC 670
QY      598 CGTGTGGGCTGGGAGGAGCATGTGCCCTAGTCAGAGGGGA 640
Db      671 CGTGTGGGCTGGGAGGAGCATGTGCCCTAGTCAGAGGGGA 713

```

Search completed: August 15, 2003, 10:21:12  
 Job time : 2938 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 09:49:08 ; Search time 83 Seconds  
(without alignments)  
231.397 Million cell updates/sec

Title: US-09-904-603-1

Perfect score: 625  
Sequence: 1 MSPDRPFKRRSFADRCKEY.....EKDEGFLYVVASQETFCF 121

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum PH seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSL/gcadata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSL/gcadata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSL/gcadata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSL/gcadata/geneseq/geneseq-emb1/AA2003.DAT:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625	100.0	121	19	AAW69979
2	625	100.0	159	23	ABP41404
3	625	100.0	199	22	ABG09793
4	625	83.4	125	22	AAW39660
5	625	83.4	125	22	AAW74775
6	625	83.4	125	22	ABG79539
7	625	83.4	130	23	ABG97779
8	625	83.4	142	22	AAW74776
9	625	83.4	155	22	AAW41446

10	521	83.4	171	21	AAW44006	Human: "CAHET" assoc
11	466.5	74.5	99	22	ABG09794	Novel human diapo
12	411	65.8	222	22	AAW06072	Novel human secret
13	359	57.4	147	21	AAW19955	Human cytoskeleton
14	301	48.2	120	22	AAW60179	Human protein SH3
15	272	43.5	54	22	ABG56081	Human liver peptid
16	272	43.5	54	22	ABG40581	Peptide #8087 origo
17	272	43.5	54	22	AAW61449	Human brain expres
18	272	43.5	54	22	AAW74229	Human bone marrow
19	272	43.5	54	22	AAW41441	Peptide #8078 origo
20	272	43.5	135	22	AAW79195	Human protein SH3
21	263	42.1	66	23	ABG32427	Human secreted pro
22	236	37.8	116	21	AAW18966	Zea mays protein 1
23	234	37.4	121	21	AAW52262	Arabidopsis thalia
24	233	37.3	151	22	AAW11284	Novel human diapo
25	230	36.8	120	21	AAW37448	Arabidopsis thalia
26	228	36.5	119	21	AAW44817	Zea mays protein 1
27	227	36.3	109	21	AAW18967	Zea mays protein 1
28	227	36.3	120	21	AAW32448	Arabidopsis thalia
29	227	36.3	120	21	AAW44107	Zea mays protein 1
30	227	36.3	120	21	AAW44172	Zea mays protein 1
31	227	36.3	120	21	AAW33761	Arabidopsis thalia
32	227	36.3	140	21	AAW32246	Arabidopsis thalia
33	226	36.2	122	21	AAW32625	Arabidopsis thalia
34	225	36.0	122	21	AAW29168	Arabidopsis thalia
35	223	35.7	115	21	AAW09473	Arabidopsis thalia
36	223	35.7	115	21	AAW51987	Arabidopsis thalia
37	220	35.2	114	21	AAW29169	Arabidopsis thalia
38	219	35.0	107	21	AAW32626	Arabidopsis thalia
39	219	35.0	122	21	AAW08604	Arabidopsis thalia
40	219	35.0	122	21	AAW41465	Arabidopsis thalia
41	219	35.0	122	21	AAW37988	Arabidopsis thalia
42	215	35.0	122	21	AAW37990	Arabidopsis thalia
43	215	35.0	184	21	AAW79987	Arabidopsis thalia
44	217	34.7	119	21	AAW10031	Arabidopsis thalia
45	217	34.7	119	21	AAW11019	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1  
AAW69979 standard; Protein: 121 AA.  
ID  
AAW69979:  
AC AAW69979:  
XX  
XX  
XX 20-NDV-1998 (first entry)  
XX  
XX Human microtubule associated protein hM3.  
DE  
XX Human: microtubule-associated protein; hMAP; cell proliferation; cancer.  
KW  
XX Homo sapiens.  
OS  
XX  
XX  
XX W09837197-A1.  
XX  
XX 27-AUG-1998.  
XX  
XX 23-FEB-1998; 98W09803744.  
XX  
XX 24-FEB-1997; 97US0805117.  
XX  
XX (INCYTE) INCYTE PHARM INC.  
XX  
XX Gol1 SK, Hillman JL:  
XX  
XX WPI: 1598-467560/40.  
XX  
XX N-PSDB: AAW43780.  
XX  
XX New isolated human microtubule-associated protein - used to develop  
XX products for use in the diagnosis, prevention and treatment of  
XX disorders associated with cell proliferation, particularly cancers

XX Claim 1; Fig 1; 58pp; English.  
PS  
XX  
CC Human microtubule-associated protein (hMAP) expression is associated with  
CC disorders of cell proliferation, in particular, cancer. The hMAP and  
CC agonists can be used to stimulate cell proliferation. Tissues or  
CC organs generated by these methods may be introduced or transplanted into  
CC a subject for the purpose of treating a variety of diseases. Antagonists  
CC or inhibitors of hMAP can be used to treat or prevent disorders  
CC associated with cell proliferation, particularly cancers. The products  
CC can also be used for detection, diagnosis and drug screening.  
XX  
SQ Sequence 121 AA;  
Query Match 100.0%; Score 625; DB 19; Length 121;  
Best Local Similarity 100.0%; Pred. No. 4.9e-65;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MPSDRPFKORRSFADRCKEVQOIRDOHPSKIPVIERKYGKQOLPVDKTKFLVDPHVM 60  
DB 1 MPSDRPFKORRSFADRCKEVQOIRDOHPSKIPVIERKYGKQOLPVDKTKFLVDPHVM 60  
QY 61 SELVKIIRRLQNLPTQAFLLVNHQSMWSVSPDIADIEQEKDEDFLYWVASQETFG 120  
DB 61 SELVKIIRRLQNLPTQAFLLVNHQSMWSVSPDIADIEQEKDEDFLYWVASQETFG 120  
QY 121 F 121  
DB 121 F 121  
Db 121 F 121  
RESULT 2  
ABP41404 ID ABP41404 standard; Protein: 169 AA.  
XX  
AC ABP41404;  
XX  
DT 22-AUG-2002 (first entry)  
XX  
DE Human ovarian antigen HFATL31, SEQ ID NO:2536.  
XX  
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
KW inflammatory condition; immune disorder; blood disorder;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disorder; urinary system disorder; drug screening;  
KW gene therapy; chromosome mapping; forensic analysis;  
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
KW antiinflammatory; gynaecological; reproductive.  
XX  
OS Homo sapiens.  
XX  
PN WO20020677-A1.  
XX  
PD 03-JAN-2002.  
XX  
PF 07-JUN-2001; 2001WO-US18569.  
XX  
PR 07-JUN-2000; 2000US-209467P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Blrse CE, Rosen CA;  
XX  
DR WPI: 2002-147878/19.  
XX  
DR N-PSDB: ABQ54481.  
XX  
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
PT useful in the prevention, treatment and diagnosis of cancer (e.g.  
PT ovarian cancer), immune disorders, cardiovascular disorders and  
PT neurological diseases -

XX Claim 11; SEQ ID NO 2536; 2922pp; English.  
PS  
XX  
CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
CC to the sequences of the invention. The invention additionally relates to  
CC recombinant vectors and host cells comprising human ovarian antigen  
CC polynucleotides, antibodies against human ovarian antigens, and the use  
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
CC treating, prognosting or preventing various ovary and/or breast-related  
CC disorders. Such conditions include ovarian cancer and breast cancer, and  
CC metastatic tumours of ovarian or breast origin, reproductive system  
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
CC vaginitis), immune disorders (e.g., congenital and acquired  
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
CC respiratory disorders, neurological disorders, gastrointestinal disorders  
CC and urinary system disorders. Ovarian antigen polypeptides and  
CC polynucleotides may also be used in screening for compounds which  
CC modulate ovarian antigen expression or activity. The polynucleotides may  
CC further be used for gene therapy, chromosome mapping, in the  
CC identification of individuals and in forensic analysis, and the  
CC polypeptides may be used as food additives or to prepare antibodies  
CC useful in disease diagnosis, drug targeting and phenotyping. The present  
CC sequence represents a human ovarian antigen of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 169 AA;  
Query Match 100.0%; Score 625; DB 23; Length 169;  
Best Local Similarity 100.0%; Pred. No. 7.6e-65;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MPSDRPFKORRSFADRCKEVQOIRDOHPSKIPVIERKYGKQOLPVDKTKFLVDPHVM 60  
DB 49 MPSDRPFKORRSFADRCKEVQOIRDOHPSKIPVIERKYGKQOLPVDKTKFLVDPHVM 108  
QY 61 SELVKIIRRLQNLPTQAFLLVNHQSMWSVSPDIADIEQEKDEDFLYWVASQETFG 120  
DB 109 SELVKIIRRLQNLPTQAFLLVNHQSMWSVSPDIADIEQEKDEDFLYWVASQETFG 168  
QY 121 F 121  
DB 169 F 169  
Db 169 F 169  
RESULT 3  
ABG09793 ID ABG09793 standard; Protein: 199 AA.  
XX  
AC ABG09793;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #9784.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX



PR	31-MAR-2000: 20000US-0540217.
PR	23-AUG-2000: 20000US-0649167.
XX	
PA	(HWS-) HXSEQ INC.
XX	
PI	Dumanac RT, Liu C, Tang YT;
XX	
DR	WPI: 2001-639362/73.
DR	N-PSDB: AAS73980.
XX	
PT	New isolated polypeptide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
XX	
PS	Claim 20: SEQ ID NO 40152; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG030377 represent novel human
CC	diagnostic amino acid sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at <a href="http://wipo.int/pub/published_pct_sequences">ftp.wipo.int/pub/published_pct_sequences</a> .
CC	
XX	
SEQ	Sequence 199 AA;
	Query Match 100.0%; Score 625; DB 22; Length 199;
	Best Local Similarity 100.0%; Pred. NO. 9.4e-65;
	Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MPSDRPKQRKRSFADCKEYVOQIRDQHPKIPVILIERKYGKQQLPVLDKTFGLVPDHYNM 60
Db	52 MPSDRPKQRKRSFADCKEYVOQIRDQHPKIPVILIERKYGKQQLPVLDKTFGLVPDHYNM 111
QY	61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPLADLYEDKEDGGLVWYASQETFG 120
Db	112 SELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPLADLYEDKEDGGLVWYASQETFG 171
QY	121 F 121
Db	172 F 172
RESULT 4	
AAAM39660	standard; Protein: 125 AA.
AC	AAAM39660:
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 2805.
XX	
KM	Human: nootropic; immunosuppressant; cytotatic; gene therapy; cancer;
KM	peripheral nervous system; neuropathy; central nervous system; CNS;
KM	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM	anyotropic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KM			leukaemia.
XX			
OS	Homo sapiens.		
XX			
PN	WO20015312-A1.		
PD			
XX	26-JUL-2001.		
FE			
XX	26-DEC-2000; 2000WO-US34263.		
PR			
XX	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
PR	29-NOV-2000; 2000US-0727344.		
XX			
PA	(HYSE-) HYSEQ INC.		
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;		
XX			
DR	WPI: 2001-442253/47.		
DR	N-PDB: AAI58816.		
PT			
PT	Novel nucleic acids and polypeptides, useful for treating disorders		
PT	such as central nervous system injuries -		
XX			
XX	Example 4; SEQ ID No 2805; 10078pp; English.		
CC			
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and		
CC	the encoded peptides (AAM58642-AAM42213) with nootropic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilisation of the activities such as: Immune system suppression,		
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic		
CC	and thrombolytic activity, cancer diagnostic and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukaemias and		
CC	C.N.S disorders.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification.		
XX			
SQ	Sequence 125 AA;		
Query Match	83.4%; Score 521; Lb 22; Length 125;		
Best Local Similarity	82.5%; Pred. No. 7.2e-53;		
Matches	9% Conservative 12; Mismatches %; Indels 0; Gaps 0;		
OY	1 MPSDRPFKORRSFADCKEYQIRDPHSKIPIVIEKYKGKQLPVLDTIKTVLHVM 60		
Dd	I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::		
OY	1 MPSEKTFKORTFGQRVEDRLIREQHFKIPVIEEKYGKKQLPVLDTIKTVLHVM 60		
Dd	61 SELVKIIIRRLQLNFQAFLLVWCHSVSSTPIADIYEDEKDEDGLYWVVASGELTG 120		
OY	I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::		
Dd	61 SELVKIIIRRLQLNANAAFLLVNGHSHVSSTPISEVESEKEDDGLFWVVASGELTG 120		
RESULT 5			
ID	AAB74775		
AC	AAB74775 standard; Protein: 125 AA.		
IT	06-JUN-2001 (first entry)		
XX			

```

DE Human MPL3 protein SEQ ID NO:4.
XX
XX Human; microtubule-associated protein 1A/1B light chain 3; MPL3;
KW hMPL3.
XX
XX Homo sapiens.
XX
XX CNI274726-A.
XX
XX 29-NOV-2000.
XX
XX 24-MAY-1999; 99CN-0106971.
XX
XX 24-MAY-1999; 99CN-0106971.
XX
XX (UYFU-) UNIV FUDAN.
XX
XX Yu L, Zhao Y, Fu Q;
XX
XX WPI; 2001-211747/22.
XX
XX N-PSDB; AAF81833.
XX
XX Human tubule related protein 1A/1B light chain 3 and its code sequence,
XX preparation and application.
XX
XX Claim 4; Page 15; 20pp; Chinese.
XX
XX The present invention describes the human microtubule-associated protein
XX 1A/1B light chain 3 (hMPL3), which is a homologue of rat MPL3. The
XX present invention also describes a method for the application and
XX production process of the hMPL3 polynucleotide and protein sequences.
XX The present sequence represents the hMPL3 protein as given in the
XX present invention.
XX
XX Sequence 125 AA:
SQ
Query Match 83.4%; Score 521; DB 22; Length 125;
Best Local Similarity 82.5%; Pred. No. 7.2e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;
OY 1 MFSDRPFKORSFADCKEVEQIRDOHPSKIPVITIERYKGEKQLPVLDTKTKFLVDPDHNM 60
DB 1 MPSEKTFKORFTEQREVEDRLIREQHPRTKIPVITIERYKGEKQLPVLDTKTKFLVDPDHNM 60
OY 61 SELVATIRRRLOLNTQAFELLVNOHSMVSVSTPIADIYEOEKDEGFLYVYASQETFG 120
DB 61 SELIKIIRRRLOLNTQAFELLVNOHSMVSVSTPISEVYSEKDEGFLYVYASQETFG 120
RESULT 6
ABG79539
ID ABG79539 standard; Protein; 125 AA.
XX
XX ABG79539;
XX
XX 15-NOV-2002 (first entry)
XX
XX Human tumour suppressor protein #25.
XX
XX Human; cancer; tumour suppressor protein; breast cancer; lymphoma;
KW adenocarcinoma; leukaemia; melanoma; myeloma; sarcoma; teratocarcinoma;
KW prostate; cervix; liver; ovary; adrenal gland; heart; brain; lung; colon;
KW placenta; skeletal muscle; synovial membrane; tonsil; kidney; uterus;
KW skin; cytosolic.
XX
XX Homo sapiens.
XX
XX WO200264775-A1.
XX
XX 22-AUG-2002.
XX
XX 12-FEB-2002; 2002WO-AU00137.
XX

```

```

PR 12-FEB-2001; 2001AU-0003052.
PR 12-FEB-2001; 2001AU-0003053.
PR 12-FEB-2001; 2001AU-0003054.
XX
XX (BION-) BIONOMICS LTD.
XX
XX Callen DF, Powell JA, Kremmidiotis G, Gardner AE, Whitmore SA;
XX
XX WPI; 2002-657597/70.
XX
XX N-PSDB; ABS64852.
XX
XX New tumour suppressor genes identified at 16q24.3, useful for
XX identifying or obtaining full-length human genes involved in the
XX tumorigenic process, or in diagnosing or treating cancer (e.g. breast
XX or cervix cancer) through gene therapy.
XX
XX Claim 22; Page 205; 218pp; English.
XX
XX The invention relates to a tumour suppressor gene and the protein it
XX encodes. The genes are useful for identifying and/or obtaining
XX full-length human genes involved in the tumorigenic process. The genes
XX the polypeptides and antibodies to the polypeptides are useful in
XX diagnosing cancer, in establishing the prognosis of a patient diagnosed
XX with cancer, or in treating cancer through gene therapy. In particular,
XX the cancer is breast cancer. The drugs that restore tumour suppressor
XX activity are useful for treating breast cancer or for manufacturing a
XX medicament for the treatment of breast cancer. The cancer may also be
XX adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma,
XX teratocarcinoma, or cancer of the prostate, cervix, liver, ovary, adrenal
XX gland, heart, brain, lung, placenta, skeletal muscle, synovial membrane,
XX tonsil, kidney, colon, uterus or skin. This sequence represents a human
XX tumour suppressor protein of the invention.
XX
XX Sequence 125 AA:
SQ
Query Match 83.4%; Score 521; DB 23; Length 125;
Best Local Similarity 82.5%; Pred. No. 7.2e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;
OY 1 MFSDRPFKORSFADCKEVEQIRDOHPSKIPVITIERYKGEKQLPVLDTKTKFLVDPDHNM 60
DB 1 MPSEKTFKORFTEQREVEDRLIREQHPRTKIPVITIERYKGEKQLPVLDTKTKFLVDPDHNM 60
OY 61 SELVATIRRRLOLNTQAFELLVNOHSMVSVSTPIADIYEOEKDEGFLYVYASQETFG 120
DB 61 SELIKIIRRRLOLNTQAFELLVNOHSMVSVSTPISEVYSEKDEGFLYVYASQETFG 120
RESULT 7
ABB97779
ID ABB97779 standard; Protein; 130 AA.
XX
XX ABB97779;
XX
XX 03-OCT-2002 (first entry)
XX
XX Human secretory polypeptide (SPTM) 31.
XX
XX Human; secretory protein; secretory polynucleotides; SPTM;
KW SPTM-related disease; somatic gene therapy; germline gene therapy;
KW severe combined immunodeficiency; intracellular parasite protection;
KW fungal parasite; protozoan parasite; cell proliferative disorder; cancer;
KW immune disorder; AIDS; neurological disorder; Parkinson's disease;
KW motor neuron disorder; demyelinating disease; multiple sclerosis;
KW meningitis; abscess; prion diseases; cerebral palsy;
KW neuroskeletal disorder; demyelinating disease; multiple sclerosis;
KW dermatomyositis; polymyositis; myopathy; myasthenia gravis;
KW mental disorder; Tourette's syndrome.
XX
XX Homo sapiens.
XX
XX WO200220756-A2.
XX

```

```

PD 14 MAR-2002.
XX
XX 30-AUG-2001: 2001WO-US27297.
XX
PR 05-SEP-2000: 2000US-229747P.
PR 05-SEP-2000: 2000US-229748P.
PR 05-SEP-2000: 2000US-229749P.
PR 05-SEP-2000: 2000US-229750P.
PR 05-SEP-2000: 2000US-229751P.
PR 05-SEP-2000: 2000US-230016P.
PR 05-SEP-2000: 2000US-230583P.
PR 06-SEP-2000: 2000US-230505P.
PR 06-SEP-2000: 2000US-230514P.
PR 06-SEP-2000: 2000US-230515P.
PR 06-SEP-2000: 2000US-230517P.
PR 06-SEP-2000: 2000US-230518P.
PR 06-SEP-2000: 2000US-230519P.
PR 06-SEP-2000: 2000US-230595P.
PR 06-SEP-2000: 2000US-230596P.
PR 06-SEP-2000: 2000US-230597P.
PR 06-SEP-2000: 2000US-230599P.
PR 06-SEP-2000: 2000US-230610P.
PR 06-SEP-2000: 2000US-230864P.
PR 06-SEP-2000: 2000US-230865P.
PR 06-SEP-2000: 2000US-230988P.
PR 06-SEP-2000: 2000US-230989P.
PR 07-SEP-2000: 2000US-230990P.
PR 07-SEP-2000: 2000US-230897P.
PR 07-SEP-2000: 2000US-230951P.
PR 07-SEP-2000: 2000US-231163P.
PR 07-SEP-2000: 2000US-231832P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
PI Momiyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
PI Gerstin EF, Peralta CH, David MH, Panzer SR, Flores V, Datto A;
PI Marwaha R, Chen AJ, Chang SC, Au AP, Iman RR;
XX
XX WPI: 2002-315658/35.
XX N-PSDB: ABL99776.
XX
XX Polynucleotide sequences encoding human secretory proteins useful for
XX gene therapy of e.g. genetic deficiency disorders, cancers, and
XX diseases caused by intracellular parasites -
XX
XX Claim 29: Page 385: 585pp: English.
XX
XX The invention comprises the amino acid and coding sequences of human
XX secretory (SPTM) proteins. The SPTM DNA and amino acid sequences are
XX useful for treating a disease or condition associated with the expression
XX of functional SPTM. The SPTM DNA sequences are useful for somatic or
XX germ-line gene therapy to correct a genetic deficiency (e.g. severe
XX combined immunodeficiency). The SPTM DNA sequences are also useful in
XX providing protection against intracellular parasites (e.g. fungal
XX parasites and protozoan parasites). The SPTM DNA and protein sequences
XX are also useful for diagnosing cell proliferative disorders, cancer,
XX immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's
XX disease), motor neuron disorders, demyelinating diseases (e.g. multiple
XX sclerosis), meningitis, abscesses, prion diseases, cerebral palsy,
XX neuroskeletal disorders, peripheral nervous system disorders,
XX dermatomycosis and polymyositis, myopathy, myasthenia gravis, and mental
XX disorders (e.g. Tourette's syndrome). Amino acid sequences ABB97749 -
XX ABB97943 represent human secretory proteins of the invention.
XX
XX Sequence 130 AA:
XX
XX Query Match 83.4%: Score 521: DH 23: Length 130:
XX Best Local Similarity 82.5%: Pred. No. 7.6e-53:
XX Matches 99: Conservative 12: Mismatches 9: Indels 0: Gaps 0:

```

```

QY 1 MPSDRPEKRSFADCKEYVQIRDQHSKIPVILIERKSGEKQLPVLDKIKELVPDHYNM 60
DB 6 MPSEKTFKQKRITFEEDKRVAVLIRKHHTKIPVILIERKSGEKQLPVLDKIKELVPDHYNM 65
QY 61 SELVKIIRRRQLNLNQAFLILVNDHWSVSTIPADIIYEDQKDEGDTLYMVVASQELIF 120
DB 66 SELIKIIRRRQLNLNQAFLILVNDHWSVSTIPSEVYESKEDGDTLYMVVASQELIF 125
RESULT 8
AAB74776
ID AAB74776 standard: protein: 142 AA.
XX
AC AAB74776:
XX
DT 06-JUN-2001 (first entry)
XX
DE Rat MPL3 protein.
XX
KW Human: microtubule-associated protein 1A/1B light chain 3; MPL3;
XX hMPL3.
XX
OS Rattus sp.
XX
PN CN1274726-A.
XX
PD 29-NOV-2000.
XX
XX 24-MAY-1999: 99CN-0106971.
XX
XX 24-MAY-1999: 99CN-0106971.
XX
XX 24-MAY-1999: 99CN-0106971.
XX
XX (UYFU-) UNIV FUJIAN.
XX
XX Yu L, Zhao Y, Fu Y;
XX
XX WPI: 2001-211747/22.
XX
XX Human tubule related protein 1A/1B light chain 3 and its cDNA sequence,
XX preparation and application -
XX
XX Example 2: Fig 1: 20pp: Chinese.
XX
XX The present invention describes the human microtubule-associated protein
XX 1A/1B light chain 3 (hMPL3), which is a homologue of rat MPL3. The
XX present invention also describes a method for the application and
XX production process of the hMPL3 polynucleotide and protein sequences.
XX The present sequence represents the rat MPL3 protein which is given
XX in comparison with the hMPL3 protein in an example from the present
XX invention.
XX
XX Sequence 142 AA:
XX
XX Query Match 83.4%: Score 521: DH 22: Length 142:
XX Best Local Similarity 82.5%: Pred. No. 8.5e-53:
XX Matches 99: Conservative 12: Mismatches 9: Indels 0: Gaps 0:
QY 1 MPSRPPKQKRSFADCKEYVQIRDQHSKIPVILIERKSGEKQLPVLDKIKELVPDHYNM 60
DB 1 MPSEKTFKQKRITFEEDKRVAVLIRKHHTKIPVILIERKSGEKQLPVLDKIKELVPDHYNM 65
QY 61 SELVKIIRRRQLNLNQAFLILVNDHWSVSTIPADIIYEDQKDEGDTLYMVVASQELIF 120
DB 61 SELIKIIRRRQLNLNQAFLILVNDHWSVSTIPSEVYESKEDGDTLYMVVASQELIF 120
RESULT 9
AAM41446
ID AAM41446 standard: protein: 155 AA.
XX
AC AAM41446:
XX
DT 22-OCT-2001 (first entry)

```









**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 10:12:08 : Search time 30 Seconds  
(without alignments)  
170.654 Million cell updates/sec

Title: US-09-904-603-1  
Perfect score: 625  
Sequence: 1 MPSDRPKQRKRSFADRCKEY.....EKDEDGLVMYASQETFGC 121

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 4210858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2-6/plodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2-6/plodata/1/1aa/5B-COMB.pep:\*  
3: /cgn2-6/plodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2-6/plodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2-6/plodata/1/1aa/PCTUS-COMB.pep:\*  
6: /cgn2-6/plodata/1/1aa/Backfilltest.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625	100.0	121	2	US-08-805-117-1
2	625	100.0	121	3	US-09-199-838-1
3	521	83.4	142	2	US-08-805-117-3
4	521	83.4	142	3	US-09-199-838-3
5	267	42.7	68	4	US-09-708-725A-4
6	263	42.1	66	4	US-09-708-725A-2
7	77.5	12.4	724	4	US-08-671-757A-9
8	70	11.2	527	3	US-08-365-822C-25
9	70	11.2	527	3	US-08-582-776C-40
10	70	11.2	527	3	US-08-434-831B-37
11	70	11.2	589	2	US-08-937-540-6
12	70	11.2	590	4	US-09-398-395A-54
13	70	11.2	590	4	US-09-887-586A-54
14	70	11.2	590	4	US-09-895-752-54
15	70	11.2	590	4	US-09-903-012B-54
16	69	11.0	704	1	US-08-190-802A-62
17	69	11.0	704	1	US-08-188-582-5
18	69	11.0	704	1	US-08-646-715-5
19	69	11.0	704	2	US-08-308-818-3
20	69	11.0	704	3	US-08-477-346-62
21	69	11.0	704	4	US-08-473-089-62
22	69	11.0	704	4	US-08-487-072A-62
23	68	10.9	251	4	US-09-107-532A-5498
24	68	10.9	566	3	US-08-335-865J-9
25	67.5	10.8	319	4	US-08-913-159-10
26	67	10.7	319	4	US-09-134-001C-5455
27	66.5	10.6	387	4	US-09-328-352-5082

28	66.5	10.6	455	4	US-09-252-991A-18265	Sequence 14265, A
29	66.5	10.6	626	3	US-09-357-251-34	Sequence 34, Appl
30	66.5	10.6	648	4	US-09-370-861A-75	Sequence 75, Appl
31	66	10.6	564	4	US-09-134-001C-2897	Sequence 2897, Ap
32	65.5	10.5	322	4	US-09-359-161-7	Sequence 161, Appl
33	65.5	10.5	3066	3	US-08-952-127-12	Sequence 12, Appl
34	65	10.4	486	3	US-08-348-518C-2	Sequence 12, Appl
35	65	10.4	872	4	US-09-107-532A-6668	Sequence 6668, Ap
36	65	10.4	1708	1	US-08-494-092-2	Sequence 2, Appl
37	65	10.4	1708	1	US-08-508-836A-2	Sequence 2, Appl
38	65	10.4	3056	1	US-08-508-836A-8	Sequence 8, Appl
39	65	10.4	3056	2	US-08-629-001A-3	Sequence 4, Appl
40	65	10.4	3056	2	US-08-874-266-2	Sequence 2, Appl
41	65	10.4	3056	3	US-08-642-274D-3	Sequence 3, Appl
42	65	10.4	3056	3	US-08-952-127-3	Sequence 3, Appl
43	65	10.4	3056	3	US-08-952-014C-3	Sequence 3, Appl
44	65	10.4	3056	4	US-09-360-416-2	Sequence 2, Appl
45	65	10.4	3057	4	US-09-360-416-3	Sequence 3, Appl

#### ALIGNMENTS

```
RESULT 1
US-08-805-117-1
: Sequence 1, Application US/08805117
: Patent No. 5955312
:
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Goli, Surya K.
: TITLE OF INVENTION: NOVEL MICROBULE-ASSOCIATED PROTEIN
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Inlyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSP for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/805,117
: FILING DATE: Bretwlich
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0211 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
:
: TELEX:
:
: INFORMATION FOR SHU ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 121 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: THYMOTO4
: CLONE: 144378
:
: US-08-805-117-1
:
: Query Match 100.0% Score 625, DB 2, Length 121:
: Best local Similarity 100.0% Pred. No. 2.2e-69:
: Matches 121: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
```

OY 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKYGKQOLPVLDTKFLVDPDHVM 60  
DB 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKYGKQOLPVLDTKFLVDPDHVM 60  
OY 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPIADIYEOEKDEGFLVMVYASQETFG 120  
DB 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPIADIYEOEKDEGFLVMVYASQETFG 120  
OY 121 F 121  
DB 121 F 121

RESULT 2  
US-09-199-838-1  
Sequence 1, Application US/09199838

Patent No. 6280733  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL MICROTUBULE-ASSOCIATED PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/199,838  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/805,117  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0211 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 121 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: THYRNOT03  
CLONE: 144378  
US-09-199-838-1

Query Match 100.0%; Score 625; DB 3; Length 121;  
Best Local Similarity 100.0%; Pred. No. 2.2e-69;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKYGKQOLPVLDTKFLVDPDHVM 60  
DB 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKYGKQOLPVLDTKFLVDPDHVM 60  
OY 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPIADIYEOEKDEGFLVMVYASQETFG 120  
DB 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPIADIYEOEKDEGFLVMVYASQETFG 120

OY 121 F 121  
DB 121 F 121

RESULT 3  
US-08-805-117-3  
Sequence 3, Application US/08805117  
Patent No. 5955312

GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL MICROTUBULE-ASSOCIATED PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/805,117  
FILING DATE: Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0211 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 142 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 455109  
US-08-805-117-3

Query Match 83.4%; Score 521; DB 2; Length 142;  
Best Local Similarity 82.5%; Pred. No. 1.7e-56;  
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKYGKQOLPVLDTKFLVDPDHVM 60  
DB 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKYGKQOLPVLDTKFLVDPDHVM 60  
OY 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPIADIYEOEKDEGFLVMVYASQETFG 120  
DB 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPISEYSESRDEGFLVMVYASQETFG 120

RESULT 4  
US-09-199-838-3  
Sequence 3, Application US/09199838  
Patent No. 6280733  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL MICROTUBULE-ASSOCIATED PROTEIN



FILED DATE: 16-AUG-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 02366.0073-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4000  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 724 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
US-08-671-757A-9

Query Match 12.4%; Score 77.5; DB 4; Length 724;  
Best Local Similarity 21.8%; Pred. No. 0.62; Mismatches 42; Indels 1; Gaps 1;  
Matches 19; Conservative 25;

QY 18 KEVQOIRDOHPKIPVIERKGEKOLPVLDK-TKFLVPDHNMSLVKIIIRRLQNP 76  
DB 533 GEVQNLDDKVKNDYPIIVGALGVAPVSLQKILKDLKHLPIKMLTLESVSDIAEV 592  
QY 77 QAFELVNOHMSVSVTPADIYEOEK 103  
DB 593 SKSFDMIIEHVASLARIAMITMYLDDK 619

RESULT 8  
US-08-369-822C-25  
Sequence 25, Application US/08369822C  
Patent No. 6015660  
GENERAL INFORMATION:  
APPLICANT: Lipkin, W. I.  
APPLICANT: Briese, Thomas  
APPLICANT: Kliche, Stefanie  
APPLICANT: Schneider, Patrick A.  
APPLICANT: Stitz, Lothar  
APPLICANT: Schneemann, Anette  
TITLE OF INVENTION: Borna Disease Viral Sequences,  
TITLE OF INVENTION: Diagnostics and Therapeutics for Central Nervous  
TITLE OF INVENTION: System Diseases  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESS: Fulbright & Jaworski, L.L.P.  
STREET: 865 South Figueroa Street, 29th Floor  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90017-2571  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: WINDOWS NT-WORDBERECT 8.0  
SOFTWARE: ASCII (DOS) TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/369,822C  
FILING DATE: 06-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Churchill, Margaret A. (Ph.D.)  
REGISTRATION NUMBER: 39,944  
REFERENCE/DOCKET NUMBER: 1279-194XX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 213/680-4518  
TELEFAX: 213/680-9200  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 527 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-369-822C-25

Query Match 11.2%; Score 70; DB 3; Length 527;  
Best Local Similarity 21.1%; Pred. No. 3.4; Mismatches 46; Indels 36; Gaps 5;  
Matches 27; Conservative 19;

QY 7 FKORSFADRCKEVOQIRD---QHPSKIPVIER---YKGEKOLPVLDKTKFLVPDHN 59  
DB 237 FESTALFGQRCNEIFGKTFEFNMHP-----VLERCTIYGDPCPADRMHROLQDHD 291  
QY 60 MSELVKIIRRLQNPLOAFLLV-----NQHSVSVSTPIA 96  
DB 292 SGIFLHNPFGIE-GYQKLMPLISMSAIHLAARVGVASAMVQDNOALAVTSRPVVA 350  
QY 97 DIVEQEKD 104  
DB 351 QYKQKN 358

RESULT 9  
US-08-582-776C-40  
Sequence 40, Application US/08582776C  
Patent No. 6077510  
GENERAL INFORMATION:  
APPLICANT: Lipkin, W. I.  
APPLICANT: Briese, Thomas  
APPLICANT: Kliche, Stefanie  
APPLICANT: Schneider, Patrick A.  
APPLICANT: Stitz, Lothar  
APPLICANT: Schneemann, Anette  
TITLE OF INVENTION: Borna Disease Viral Sequences,  
TITLE OF INVENTION: Diagnostics and Therapeutics for Central Nervous  
TITLE OF INVENTION: System Diseases  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESS: Fulbright & Jaworski, L.L.P.  
STREET: 865 South Figueroa Street, 29th Floor  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90017-2576  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: WINDOWS NT  
SOFTWARE: ASCII DOS TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/582,776C  
FILING DATE: 04-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/369,822  
FILING DATE: 06-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/434,831  
FILING DATE: 04-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Churchill, Margaret A.  
REGISTRATION NUMBER: 39,944  
REFERENCE/DOCKET NUMBER: 1279-194C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 213/680-4518  
TELEFAX: 213/680-9200  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 527 amino acids  
TYPE: amino acid  
TOPOLOGY: linear



Db 182 DTKG 185

## RESULT 12

US-09-398-395A-54  
; Sequence 54, Application US/09398395A  
; Patent No. 6468772  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 6468772, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/398,395A  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/100,993  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 54  
; LENGTH: 590  
; TYPE: PR1  
; ORGANISM: Salvia officinalis  
US-09-398-395A-54

## Query Match

11.2%; Score 70; DB 4; Length 590;  
Best Local Similarity 22.6%; Pred. No. 4;  
Matches 28; Conservative 26; Mismatches 50; Indels 20; Gaps 4;

QY 6 PFQORSFADRCREVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPDHVNMSELYK 65

Db 73 PYKQORHFNQALIMQVRLKVKMEAI-----QQLLEIDDLQYLGLSYFFODEIKQ 125

QY 66 IIR-----RRLQNPQAFLLVNOHSMVSVSTPIADIYOEKEDGFLYMYVASQ 116

Db 126 ILSIHNEPRYFNNLDLYFTALGFRLIRQHG-FNVSEVDYDCKRIKCSDF--FNANLAQ 181

QY 117 EFTFG 120

Db 182 DTKG 185

## RESULT 13

US-09-887-586A-54  
; Sequence 54, Application US/09887586A  
; Patent No. 6495354  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 6495354, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/887,586A  
; PRIOR FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 54  
; LENGTH: 590  
; TYPE: PR1  
; ORGANISM: Salvia officinalis

US-09-887-586A-54

## Query Match

11.2%; Score 70; DB 4; Length 590;  
Best Local Similarity 22.6%; Pred. No. 4;  
Matches 28; Conservative 26; Mismatches 50; Indels 20; Gaps 4;

QY 6 PFQORSFADRCREVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPDHVNMSELYK 65

Db 73 PYKQORHFNQALIMQVRLKVKMEAI-----QQLLEIDDLQYLGLSYFFODEIKQ 125

QY 66 IIR-----RRLQNPQAFLLVNOHSMVSVSTPIADIYOEKEDGFLYMYVASQ 116

Db 126 ILSIHNEPRYFNNLDLYFTALGFRLIRQHG-FNVSEVDYDCKRIKCSDF--FNANLAQ 181

QY 117 EFTFG 120

Db 182 DTKG 185

## RESULT 14

US-09-895-752-54  
; Sequence 54, Application US/09895752  
; Patent No. 6559297  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 6559297, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/895,752  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/100,993  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 54  
; LENGTH: 590  
; TYPE: PR1  
; ORGANISM: Salvia officinalis  
US-09-895-752-54

## Query Match

11.2%; Score 70; DB 4; Length 590;  
Best Local Similarity 22.6%; Pred. No. 4;  
Matches 28; Conservative 26; Mismatches 50; Indels 20; Gaps 4;

QY 6 PFQORSFADRCREVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPDHVNMSELYK 65

Db 73 PYKQORHFNQALIMQVRLKVKMEAI-----QQLLEIDDLQYLGLSYFFODEIKQ 125

QY 66 IIR-----RRLQNPQAFLLVNOHSMVSVSTPIADIYOEKEDGFLYMYVASQ 116

Db 126 ILSIHNEPRYFNNLDLYFTALGFRLIRQHG-FNVSEVDYDCKRIKCSDF--FNANLAQ 181

QY 117 EFTFG 120

Db 182 DTKG 185

## RESULT 15

US-09-903-012B-54  
; Sequence 54, Application US/09903012B  
; Patent No. 6569656  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 6569656, Joseph P.  
; APPLICANT: Starks, Courtney M.



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 10:15:34 : Search time 53 seconds

(Without alignments)  
290.617 Million cell updates/sec

Title: US-09-904-603-1

Perfect score: 625

Sequence: 1 MSPDRPFKORRSPADRCREY.....EKDEDGLYVVASQETFCF 121

Scoring table: H10SUM62

Gapop 10.0 , Gapext 0.5

Searched: 479057 seqs, 127295195 residues

Total number of hits satisfying chosen parameters: 479057

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*

5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_NEW\_PUBCOMB.pep:\*

7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*

9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*

10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*

11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep:\*

12: /cgn2\_6/ptodata/1/pubppaa/US09C\_NEW\_PUB.pep:\*

13: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep:\*

14: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep:\*

15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep:\*

16: /cgn2\_6/ptodata/1/pubppaa/US10C\_NEW\_PUB.pep:\*

17: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*

18: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625	100.0	121	11	US-09-904-603-1
2	521	83.4	142	11	US-09-904-603-3
3	521	83.4	142	12	US-10-205-194-71
4	521	83.4	171	9	US-09-925-301-1051
5	272	43.5	53	9	US-09-864-761-44587
6	267	42.7	68	15	US-10-277-156-4
7	263	42.1	66	15	US-10-277-156-2
8	248	39.7	117	15	US-10-205-342-23
9	71	11.4	501	9	US-09-815-242-5292
10	71	11.4	504	9	US-09-815-242-12503
11	70	11.2	590	9	US-09-887-585A-54
12	70	11.2	590	9	US-09-903-012-54
13	70	11.2	590	11	US-09-900-797-54
14	69.5	11.1	1233	10	US-09-738-626-4312
15	68.5	11.0	362	15	US-10-177-293-375

16	68.5	11.0	458	10	US-09-764-864-1164
17	68.5	11.0	578	15	US-10-137-418A-3
18	68.5	11.0	1360	15	US-10-137-418A-2
19	68	10.9	326	14	US-10-057-989-1
20	68	10.9	1616	11	US-09-820-843A-16
21	67	10.7	1236	15	US-10-128-714-3224
22	65	10.4	497	10	US-09-902-525-32
23	65	10.4	1814	9	US-09-920-552-103
24	65	10.4	3056	8	US-08-984-090-2
25	65	10.4	3036	11	US-09-825-476-3
26	65	10.4	3056	15	US-10-307-077-1
27	65	10.4	3056	15	US-10-175-225-2
28	65	10.4	3056	15	US-10-059-585-41
29	65	10.4	3057	15	US-10-175-225-3
30	64.5	10.3	324	10	US-09-981-876-181
31	64.5	10.3	324	11	US-09-148-545-181
32	64	10.2	317	15	US-10-023-282-504
33	64	10.2	1008	10	US-09-866-582-39
34	63.5	10.2	83	9	US-09-867-550-498
35	63	10.1	220	9	US-09-815-242-10585
36	63	10.1	1718	15	US-10-223-070-21
37	62.5	10.0	190	14	US-10-108-605-167
38	62.5	10.0	267	15	US-10-189-346-89
39	62.5	10.0	283	14	US-10-042-417-50
40	62.5	10.0	586	11	US-09-878-131-2
41	62.5	10.0	1257	15	US-10-128-714-8224
42	62.5	10.0	26926	10	US-09-759-508B-2
43	62	9.9	220	9	US-09-938-803-4
44	62	9.9	335	15	US-10-156-761-7713
45	62	9.9	384	10	US-09-801-876B-6

#### ALIGNMENTS

RESULT 1

US-09-904-603-1

Sequence 1, Application US/09904603

Publication No. US20030059612A1

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

Goll, Surya K.

TITLE OF INVENTION: NOVEL MICROTUBULE ASSOCIATED PROTEIN

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESS: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/904,603

FILING DATE: 12-Jul-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/805,117

FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0211 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4196

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

Sequence 1164, Ap

Sequence 3, Appl1

Sequence 2, Appl1

Sequence 1, Appl1

Sequence 15, Appl1

Sequence 1424, Ap

Sequence 32, Appl1

Sequence 104, Ap

Sequence 2, Appl1

Sequence 3, Appl1

Sequence 1, Appl1

Sequence 4, Appl1

Sequence 41, Appl1

Sequence 3, Appl1

Sequence 181, Ap

Sequence 181, Ap

Sequence 504, Ap

Sequence 498, Ap

Sequence 10585, A

Sequence 21, Appl1

Sequence 167, Ap

Sequence 89, Appl1

Sequence 50, Appl1

Sequence 2, Appl1

Sequence 824, Ap

Sequence 2, Appl1

Sequence 771, Ap

Sequence 6, Appl1

SEQUENCE DESCRIPTION: SEQ ID NO: 3

; LENGTH: 171

; TYPE: PRT



```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Human
US-10-277-156-2

Query Match
Best Local Similarity 42.1%; Score 263; DB 15; Length 66;
Matches 52; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 60 MSELVKIIRRLQINPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLYVNASQETP 119
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MSELIKIIRRLQINANQAFLLVNGHSMVSVSTPISEYSEKDEGFLYVNASQETP 60

QY 120 G 120
Db 61 G 61

RESULT 8
US-10-205-342-23
; Sequence 23, Application US/10205342
; Publication No. US20030108906A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alstair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Plincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: W-1-A-018198
; CURRENT APPLICATION NUMBER: US/10/205,342
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Protein: Ganglioside expression factor 2
US-10-205-342-23

Query Match
Best Local Similarity 39.7%; Score 248; DB 15; Length 117;
Matches 48; Conservative 29; Mismatches 36; Indels 2; Gaps 2;

QY 7 FKQRRSFADRCKEVQOIRDOHPSKIPYIERKYGKQQLPYLDKTKRFLVPHVNMSELVKI 66
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 5 FKEDHSLEHRCVESAKIRAKYPRDPRVIVEKVSQ-SQIVIDKRRKLVPSDITVAQGMWI 63

QY 67 IRRRLQINPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLYVNASQETP 121
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 64 IRRRIQLPSKRALFLFVDR-TVPOSSLTQMQLYEKEDGFLYVNASGENTGCF 117

RESULT 9
US-09-815-242-5292
; Sequence 5292, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlisen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
```

```

; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5292
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(501)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-5292

Query Match
Best Local Similarity 11.4%; Score 71; DB 9; Length 501;
Matches 26; Conservative 32; Mismatches 47; Indels 26; Gaps 6;

QY 8 KORRSFADRCKEVQOIRDOHPSKIPYIERKYGKQQLPYLDKTKRFLVPHD 57
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 231 KNRMEETVYKDDTTIDEEFTDYGKORPSSEIPLTDNNENKALASLDLTNSITSNS 290

QY 58 VNMSLVKIIIR-RLQINPTQAFLLVNOHSMVSVST-----PIADIYEQEKDEGFL 108
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 291 SQKKTVDKVVKKPKDQSTGKKEF--VIKEPTKPSIPYTEIRTKKPPSPSKOELKEK--- 345

QY 109 LYNVYASQETP 119
Db 346 --IYSSSPORF 354

RESULT 10
US-09-815-242-12503
; Sequence 12503, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlisen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
```

```

: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO: 12503
: LENGTH: 504
: TYPE: PRI
: ORGANISM: Staphylococcus aureus
US-09-815-242-12503

Query Match          11.4%  Score 71;  DB 9;  Length 504;
Best Local Similarity 19.8%;  Pred. No. 9.6;
Matches 26;  Conservative 32;  Mismatches 47;  Indels 26;  Gaps 6;

UY      8  KORSEFADCKEVOQIRD-----OHPSKIPVILIERYKGEKOL--PVLDTKFLVPDH 57
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      249  KNRMEFEITYKDELTINDEFDTGKORPSFIPTLTNNENKALASRLTISITSNS 308
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
UY      58  VNNSLKVLIIR-RLQINPTQAFLLVNOHSMVSVT-----PIADIYEDKEDGDF 108
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      109  SOKTVDKVKKPKDQSTGNKEF--VIKEPTKPSIPTYEITKPKMPSIKQETIKER--- 363
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
UY      109  LYWYASQETP 119
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      364  --IIYSSPDRP 372
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 11
US-09-887-586A-54
: Sequence 54, Application US/09887586A
: Patent No. US20020094556A1
: GENERAL INFORMATION:
: APPLICANT: Chappell, Joseph
: APPLICANT: No. US20020094556A11, Joseph P.
: APPLICANT: Starks, Courtney M.
: APPLICANT: Manna, Kathleen R.
: TITLE OF INVENTION: SYNTHASES
: FILE REFERENCE: 07678-025001
: CURRENT APPLICATION NUMBER: US/09/887,586A
: PRIOR FILING DATE: 2001-06-22
: PRIOR APPLICATION NUMBER: 09/398,395
: PRIOR FILING DATE: 1999-09-17
: PRIOR APPLICATION NUMBER: 60/130,628
: PRIOR FILING DATE: 1999-04-22
: PRIOR APPLICATION NUMBER: 60/150,262
: PRIOR FILING DATE: 1999-08-23
: NUMBER OF SEQ ID NOS: 58
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 54
: LENGTH: 590
: TYPE: PRI
: ORGANISM: Salvia officinalis
US-09-887-586A-54

Query Match          11.2%  Score 70;  DB 9;  Length 590;
Best Local Similarity 22.6%;  Pred. No. 16;
Matches 28;  Conservative 26;  Mismatches 50;  Indels 20;  Gaps 4;

UY      6  PFKORSEFADCKEVOQIRQOHPSKIPVILIERYKGEKOLPVLDTKFLVPDHVMSLKV 65
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      73  PYKQHRHFNQAEILIMQVRLMLKVKMKA1-----QOLELIDLOYLGLSYFFUDETKO 125
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
UY      66  IIR-----RRLQINPTQAFLLVNOHSMVSVTPIADIYEDKEDGDFLYWYASQ 116
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      126  ILSSTHNEPRYFHNNDLYFTALGFRILRQHG- NVSEDEVDFCEKIERCSD---FNANLAO 181
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
UY      117  EFTG 120
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      182  DTKG 185
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

```

RESULT 12
US-09-903-012-54
: Sequence 54, Application US/09903012
: Patent No. US20020094557A1
: GENERAL INFORMATION:
: APPLICANT: Chappell, Joseph
: APPLICANT: No. US20020094557A11, Joseph P.
: APPLICANT: Starks, Courtney M.
: APPLICANT: Manna, Kathleen R.
: TITLE OF INVENTION: SYNTHASES
: FILE REFERENCE: 07678-025001
: CURRENT APPLICATION NUMBER: US/09/903,012
: PRIOR FILING DATE: 2001-07-11
: PRIOR APPLICATION NUMBER: 09/448,495
: PRIOR FILING DATE: 1999-09-17
: PRIOR APPLICATION NUMBER: 60/100,993
: PRIOR FILING DATE: 1998-09-18
: PRIOR APPLICATION NUMBER: 60/140,628
: PRIOR FILING DATE: 1999-04-22
: PRIOR APPLICATION NUMBER: 60/150,262
: PRIOR FILING DATE: 1999-08-23
: NUMBER OF SEQ ID NOS: 58
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 54
: LENGTH: 590
: TYPE: PRI
: ORGANISM: Salvia officinalis
US-09-903-012-54

Query Match          11.2%  Score 70;  DB 9;  Length 590;
Best Local Similarity 22.6%;  Pred. No. 16;
Matches 28;  Conservative 26;  Mismatches 50;  Indels 20;  Gaps 4;

UY      6  PFKORSEFADCKEVOQIRQOHPSKIPVILIERYKGEKOLPVLDTKFLVPDHVMSLKV 65
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      73  PYKQHRHFNQAEILIMQVRLMLKVKMKA1-----QOLELIDLOYLGLSYFFUDETKO 125
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
UY      66  IIR-----RRLQINPTQAFLLVNOHSMVSVTPIADIYEDKEDGDFLYWYASQ 116
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      126  ILSSTHNEPRYFHNNDLYFTALGFRILRQHG- NVSHVDFCEKIERCSD---FNANLAO 181
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
UY      117  EFTG 120
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      182  DTKG 185
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 13
US-09-900-797-54
: Sequence 54, Application US/09900797
: Publication No. US20030087406A1
: GENERAL INFORMATION:
: APPLICANT: Chappell, Joseph
: APPLICANT: No. US20030087406A11, Joseph P.
: APPLICANT: Starks, Courtney M.
: APPLICANT: Manna, Kathleen R.
: TITLE OF INVENTION: SYNTHASES
: FILE REFERENCE: 07678-025001
: CURRENT APPLICATION NUMBER: US/09/900,797
: PRIOR FILING DATE: 2001-07-06
: PRIOR APPLICATION NUMBER: 09/398,395
: PRIOR FILING DATE: 1999-09-17
: PRIOR APPLICATION NUMBER: 60/130,628
: PRIOR FILING DATE: 1999-04-22
: PRIOR APPLICATION NUMBER: 60/150,262
: PRIOR FILING DATE: 1999-08-23
: NUMBER OF SEQ ID NOS: 58
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 54
: LENGTH: 590
: TYPE: PRI
: ORGANISM: Salvia officinalis
```

US-09-900-797-54

Query Match 11.2%; Score 70; DB 11; Length 590;  
Best Local Similarity 22.6%; Pred. No. 16;  
Matches 28; Conservative 26; Mismatches 50; Indels 20; Gaps 4;

OY 6 PFKORRSFADCKEVOQIRDOHPSKIPYIERYKGEKOLPYLDKTRFLVPHVNSSELYK 65  
DB 73 PYKEORHFNROAEILMQLVYMLKVMKMEAI-----QOLELIDDLQYLGISYFQDEIKQ 125  
OY 66 IIR-----RRLQNPQAFLLVNOHSMVSTPIADIYDEKEDGFLVNVYSQ 116  
DB 126 ILSIHNEPRYFHNNDLYFTALGFRLKHGP-NVSEDFQCFKIEKCD--FNANLAQ 181  
OY 117 ETPG 120  
DB 182 DTG 185

RESULT 14

US-09-738-626-4312  
Sequence 4312, Application US/09738626  
Publication No. US20020197605A1

GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAOKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 4312  
LENGTH: 1233  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4312

Query Match 11.1%; Score 69.5; DB 10; Length 1233;  
Best Local Similarity 18.8%; Pred. No. 48;  
Matches 34; Conservative 19; Mismatches 39; Indels 89; Gaps 5;

OY 1 MDSORPKQRSFADCKEVOQIRDO----- 26  
DB 1 MTSRRLKQERSFADDLQDLTKLNDQLFTNAKLQARISGNDGKTRTPPLALDQ 60  
OY 27 -----HPSKIVIERKGEKOLPYLDKTRFL----- 53  
DB 61 LTVBEYETIATLVEAVGNSKPAIDKDLFEYPLVFLALSGTAMLDAQEGFWPAFWK 120  
OY 54 -----VDPHVN-----MSELVKIIRRRLONPQAFLL-----LVNOSHMSVSTPI 95  
DB 121 RTQVSVEPHYDAIRKELVNSIRK-----NGLETFSLADLNREYVGLQHSGLSADML 176  
OY 96 A 96  
DB 177 A 177

RESULT 15

US-10-177-293-375  
Sequence 375, Application US/10177293  
Publication No. US20030124128A1

GENERAL INFORMATION:  
APPLICANT: Lillie, James  
APPLICANT: Glat, Karen  
APPLICANT: Zhao, Xumei  
APPLICANT: Gannavarpu, Manjula  
APPLICANT: Kamatkar, Shubhangi  
APPLICANT: Mertens, Maureen  
APPLICANT: Myer, Vic  
APPLICANT: Wang, Youzhen  
APPLICANT: Xu, Yongyao  
APPLICANT: Hoersch, Sebastian  
APPLICANT: Monahan, John  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Bast Jr., Robert C.  
APPLICANT: Hortobagyi, Gabriel N.  
APPLICANT: Pusztai, Lajos  
APPLICANT: Meric, Funda  
APPLICANT: Sahlin, Aysegul  
APPLICANT: Mills, Gordon B.  
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT  
FILE REFERENCE: MRI-038  
CURRENT APPLICATION NUMBER: US/10/177,293  
CURRENT FILING DATE: 2002-06-21  
PRIOR APPLICATION NUMBER: US 60/299,887  
PRIOR FILING DATE: 2001-06-21  
PRIOR APPLICATION NUMBER: US 60/301,572  
PRIOR FILING DATE: 2001-06-27  
PRIOR APPLICATION NUMBER: US 60/306,501  
PRIOR FILING DATE: 2001-07-18  
PRIOR APPLICATION NUMBER: US 60/325,002  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: US 60/362,585  
PRIOR FILING DATE: 2002-03-05  
PRIOR APPLICATION NUMBER: US 60/xxx,xxx  
PRIOR FILING DATE: 2002-05-14  
NUMBER OF SEQ ID NOS: 506  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 375  
LENGTH: 362  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-177-293-375

Query Match 11.0%; Score 68.5; DB 15; Length 362;  
Best Local Similarity 29.3%; Pred. No. 12;  
Matches 24; Conservative 12; Mismatches 37; Indels 9; Gaps 3;

OY 41 EKOLPYLDK-TKFLVPHVNSSELYKTIIRRLQNPQAFLLVNOHSMVSTPIADIY 99  
DB 285 EKALAILSQPTPSLVVHERLKNLKTIVVKSQ---NYNIFOLNLYAVVISQC-----IY 336  
OY 100 EOKEDEGFLVNVASQETFGF 121  
DB 337 RHRKDHKTSILOKMEQEVENF 358

Search completed: August 11, 2003, 10:24:37  
Job time : 53 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: August 11, 2003, 10:09:23 ; Search time 38 Seconds

(without alignments)  
306.221 Million cell updates/sec

Title: US-09-904-603-1

Perfect score: 625

Sequence: 1 MPSDRPKORRSFADCKEY.....EKDEDGLYMYVASQETFGF 121

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

# SUMMARIES

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	521	83.4	142	2	A53624
2	381.5	61.0	130	2	T27920
3	227	36.3	120	2	H84470
4	225	36.0	122	2	C84887
5	224	35.8	121	2	T40818
6	219	35.0	107	2	T49105
7	219	35.0	122	2	C85058
8	217	34.7	121	2	T47875
9	216	34.6	117	2	S45432
10	203	32.5	117	2	JC7698
11	201.5	32.2	149	2	T02148
12	194	31.0	123	2	T15740
13	140	22.4	165	2	B71432
14	85	13.6	841	2	B71212
15	85	13.6	1210	2	H88451
16	79.5	12.7	282	2	T27449
17	77.5	12.4	500	2	E96739
18	77.5	12.4	724	2	A49217
19	77.5	12.4	724	2	C81361
20	77.5	12.3	627	2	T49837
21	75	12.0	592	2	T50154
22	74.5	11.9	660	2	A81704
23	74	11.8	790	2	T19683
24	73	11.7	746	2	S60480
25	73	11.7	746	2	T43648
26	72.5	11.6	224	2	F69337
27	72.5	11.6	682	2	T23813
28	72	11.5	483	2	F64204
29	72	11.5	1046	1	GNMVC

## ALIGNMENTS

### RESULT 1

A53624 microtubule-associated protein 1 light chain 3 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 24 Sep-1995

C:Accession: A53624; S45180

R:Mann, S.S.; Hamada, J.A.

J. Biol. Chem. 269, 11492-11497, 1994

A:Title: Molecular characterization of light chain 3. A microtubule binding subunit of

A:Reference number: A53624; M0104209341; PMID:7908909

A:Accession: A53624

A:Molecule type: mRNA

A:Residues: 1-142 <MAN>

A:Cross-references: EMBL:005764; NID:0455108; P1DN:AAA20645.1; F11:0455109

C:Superfamily: hypothetical protein YH078c

Query Match 83.4% Score 521 DB 2 Length 142

Best Local Similarity 82.5% Prod. No. 6,7e+44

Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0

OY 1 MPSDRPKORRSFADCKEYVQIIRQNRKIVVLRKYKQYRVDKTKTVRNNM 50  
 DB 1 MSESEKTKORRSFADCKEYVQIIRQNRKIVVLRKYKQYRVDKTKTVRNNM 60  
 OY 61 SEVKTIRRRKQINPTQAFELLVNQMYSVSTRADIVYKDEKDEHFIYVYASQETFG 120  
 DB 61 SEVKTIRRRKQINPTQAFELLVNQMYSVSTRADIVYKDEKDEHFIYVYASQETFG 120

### RESULT 2

T27920 hypothetical protein ZK593.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000

C:Accession: T27920

R:McMurray, A.

submitted to the EMBL Data Library, February 1996

A:Reference number: Z20440

A:Accession: T27920

A>Status: preliminary; translated from GH/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-130 <WIL>

A:Cross-references: EMBL:Z69485; P1DN:AAV4421.1; USFDB:GN00022; CESP:ZK593.6

A:Experimental source: clone ZK593

C:Genetics:

A:Gene: CESP:ZK593.6

A:Map position: 4

A:Introns: 24/1: 80/2

C:Superfamily: hypothetical protein YH078c

Query Match 61.0% Score 381.5 DB 2 Length 140

Best Local Similarity 58.3% Prod. No. 2.8e+30

	Matches	70;	Conservative	27;	Mismatches	20;	Indels	3;	Gaps	1;
OY	1	MSDFRPFKORRSFADRCKEVQOIRDOHPSKIPIYIERKYGEKQLPYLDKTKFLVPDHNMSLTVK	60							
Db	14	VPS---FKRRPPEHEGOKNVEEIRSQOPAKPKVPIIERYEDGERSLPLMDRCFLVEPHITV	70							
OY	61	SELVKTIIRRLQNLPTQAFFLLYNQHSWVSFTPADIEDEKDDGLFYVYASQETFG	120							
Db	71	AELMSTIVRRRLQLHPQOAFELLVNERSMVSNMSMTLSQERDPDGFVVNYVTSPAFG	130							
RESULT 3	H84470	probable microtubule-associated protein [imported] - Arabidopsis thaliana								
C:Species:	Arabidopsis thaliana (mouse-ear cress)									
C>Date:	02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001									
C:Accession:	H84470									
R:Lili, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;										
Kuss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,										
Nature 402:761-768, 1999										
A:Title:	Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.									
A:Reference number:	A84420; MIDB:20083487; PMID:10617197									
A:Status:	Preliminary									
A:Molecule type:	DNA									
A:Residues:	1-120 <STO>									
A:Cross-references:	GB:AB002093; NID:g4581162; PIDN:AAD24645.1; GSPDB:GN00139									
C:Genetics:										
A:Gene:	AT2g05630									
A:Map position:	2									
C:Superfamily:	hypothetical protein YBL078C									
Query Match	36.3%; Score 227; DB 2; Length 120;									
Best Local Similarity	37.7%; Pred. No. 3,5e-15;									
Matches	43; Conservative 32; Mismatches 37; Indels 2; Gaps 2;									
OY	7	EKORRSFADRCKEVQOIRDOHPSKIPIYIERKYGEKQLPYLDKTKFLVPDHNMSLTVK	66							
Db	6	FKHHPLEKRQAEARIREKYPDIRIVERAE-KSDPIDDKKKLVLPADLVGGFVYV	64							
OY	67	IRRRLQNLPTQAFFLLYNQHSWVSFTPADIEDEKDDGLFYVYASQETFG	120							
Db	65	VKKIKLSPEKAIFTFV-KILPPPTALINSAITYEHKDEDGFLYMGSGENTFG	117							
RESULT 4	C84887	probable microtubule-associated protein [imported] - Arabidopsis thaliana								
C:Species:	Arabidopsis thaliana (mouse-ear cress)									
C>Date:	02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001									
C:Accession:	C84887									
R:Lili, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;										
Kuss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,										
Nature 402:761-768, 1999										
A:Title:	Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.									
A:Reference number:	A84420; MIDB:20083487; PMID:10617197									
A:Status:	Preliminary									
A:Molecule type:	DNA									
A:Residues:	1-122 <STO>									
A:Cross-references:	GB:AB002093; NID:g6598366; PIDN:AAF18602.1; GSPDB:GN00139									
C:Genetics:										
A:Gene:	AT2g45170									
A:Map position:	2									
C:Superfamily:	hypothetical protein YBL078C									
Query Match	36.0%; Score 225; DB 2; Length 122;									
Best Local Similarity	39.5%; Pred. No. 5,5e-15;									
Matches	47; Conservative 29; Mismatches 31; Indels 12; Gaps 4;									
OY	7	EKORRSFADRCKEVQOIRDOHPSKIPIYIERKYGEKQLPYLDKTKFLVPDHNMSLTVK	65							



Db 62 IFVHV-KNLTPTALMSATIEFHKDEGFLYMTSGENTFG 102

## RESULT 7

C85058

Probable symbiosis-related protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001

C:Accession: C85058

R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402: 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: C85058

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-122 <STD>

A:Cross-references: GB:NC\_001268; NID:q7267220; PIDN:CA80827.1; GSPDB:GN00140

C:Genetics:

A:Gene: At4g04620

C:Superfamily: hypothetical protein YBL078c

## Query Match

Best Local Similarity 35.0%; Score 219; DB 2; Length 122;

Matches 42; Conservative 31; Mismatches 39; Indels 2; Gaps 2;

Db 7 FKORRSFADCKEVOQIRDOHPSKIPVIERKGEKQLPVLDKTRFLVPHVNSLVKI 66

Db 6 FKLSNPLEMRMAESTRIKAKYPERVPIVER-ACQSHVPOIDKKRYLVPAIDLIGQVYV 64

Db 67 IRRRLQNPQAFLLVNOHSMVSVTPIDVYEQFDEDFLYMYVASQETFG 120

Db 65 VKRKIKGAEKAIIVFV-KNLTPTALMSATIEFHKDEGFLYMTSGENTFG 117

## RESULT 8

T47875

hypothetical protein T4C21.50 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 21-Jul-2000

C:Accession: T47875

R:Choiste, N.; Brodtier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa-

W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24479

A:Accession: T47875

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-121 <CHO>

A:Cross-references: EMBL:AL162295

A:Experimental source: cultivar Columbia; BAC clone T4C21

C:Genetics:

A:Map position: 3

A:Initrons: 13/1; 31/3; 49/2; 89/1

A:Note: T4C21.50

C:Superfamily: hypothetical protein YBL078c

Query Match 34.7%; Score 217; DB 2; Length 121;

Best Local Similarity 37.0%; Pred. No. 3.3e-14;

Matches 44; Conservative 32; Mismatches 39; Indels 4; Gaps 3;

Db 3 SDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKGEKQLPVLDKTRFLVPHVNS 61

Db 2 SNVSFKODHDEPKKAKALRIREKYSRPVIVE-KSEKSDIPNDKKTLVPAIDLIG 59

Db 62 ELVKIIRRLQNPQAFLLVNOHSMVSVTPIDVYEQFDEDFLYMYVASQETFG 120

Db 60 QFVYVIRKRIQLSSEKAFIFVVD-NVLPITGAMMSTIYDEKKEGDFLYMTSGENTFG 117

## RESULT 9

S45432

hypothetical protein YHL078c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YBL0732

C:Species: Saccharomyces cerevisiae

C:Date: 09-Jun-1994 #sequence\_revision 09-Sep-1994 #text\_change 19-Apr-2002

C:Accession: S45432; S45814; S45819; S59229

R:Obermaler, R.; Gassenhuber, J.; Piravandi, E.; Domdey, H.

submitted to the EMBL Data Library, May 1994

A:Description: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces

A:Reference number: S45387

A:Accession: S45432

A:Molecule type: DNA

A:Residues: 1-117 <OBE>

A:Cross-references: EMBL:X79489; NID:q496661; PIDN:CA456032.1; PID:q49795

A:Experimental source: Strain S288c

R:Contreras, R.; Fiers, W.; Lodghe, M.; Molmans, F.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45802

A:Accession: S45814

A:Molecule type: DNA

A:Residues: 1-117 <CON>

A:Cross-references: EMBL:Z35839; NID:q536124; PIDN:CA484899.1; PID:q536125; MIPS:YBL0

A:Experimental source: Strain S288c

R:Domdey, H.; Gassenhuber, H.; Obermaler, R.; Piravandi, E.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45815

A:Accession: S45819

A:Molecule type: DNA

A:Residues: 1-117 <DM>

A:Cross-references: EMBL:Z35839; NID:q536124; PIDN:CA484899.1; PID:q536125; MIPS:YBL0

A:Experimental source: Strain S288c

R:Obermaler, R.; Gassenhuber, J.; Piravandi, E.; Domdey, H.

Yeast 11, 1103-1112, 1995

A:Title: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces cere

A:Reference number: S59184; MUID:96076635; PMID:7502586

A:Accession: S59229

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-117 <OHV>

A:Cross-references: EMBL:X79489; NID:q496661; PIDN:CA456032.1; PID:q49795

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1994

C:Genetics:

A:Gene: SGD:A077

A:Cross-references: SCD:S0000174

A:Map position: 2L

C:Superfamily: hypothetical protein YBL078c

Query Match 34.6%; Score 215; DB 2; Length 117;

Best Local Similarity 37.8%; Pred. No. 4e-14;

Matches 45; Conservative 32; Mismatches 32; Indels 10; Gaps 4;

Db 3 SDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKGEKQLPVLDKTRFLVPHVNS 61

Db 7 SEYFPEKKA-----ESERIALDFKKNIPVIGF-RAEKSDIPEDIKKTLVPAIDLIG 58

Db 62 ELVKIIRRLQNPQAFLLVNOHSMVSVTPIDVYEQFDEDFLYMYVASQETFG 120

Db 59 QFVYVIRKRIQLSSEKAFIFVVD-TLPPTALMSATIEFHKDEGFLYMTSGENTFG 116

## RESULT 10

JC7698

gecl protein - quinea piz

C:Species: Cavia porcellus (guinea pig)

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 22-Nov-2001

C:Accession: JC7698

R:Venier-Maguin, S.; Muller, S.; Saljot, M.; Radom, J.; Musard, J.F.; Adami, P.; Dul

Biochem. Biophys. Res. Commun. 284, 118-125, 2001

A:Title: A novel early estrogen-regulated gene gecl encodes a protein related to GABA

A:Reference number: JC7698; MUID:21266536; PMID:11374880

A:Accession: JC7698

A:Molecule type: mRNA

A:Residues: 1-117 <VEK>



A:Experimental source: strain 073  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH1962

Query Match	13.6%	Score 85	DB 2	Length 841
Best Local Similarity	26.6%	Pred. No. 2.8		
Matches	29	Conservative	21	Mismatches 39
				Indels 20
				Gaps 4

QY 8 KÖRSFADREKVEODIRDOHPSKIPIV---TERKYCKÖLVLDTKIFIVPHVNMSLV 64  
 II : : : : : : : : : : : : : : : :  
 Db 248 KÖLSFSSESDIEDIOIDMLPAHAHVMEELKKERPIVDYLK-QFIYNODEIKEYD 306

```

QY      65 KIHKRLQNPQTQAEFLVNQHSMSVSTPIADIYEKEDKEDGFLTWY 113
      ||      | : : | : | : | : | : | : | : | : | : | : |
Db      107 DIIN-----FIKQHNHMFISNGPYIIDVVPEN-----LYLKY 339

```

```

RESULT 15
H88451
protein ZC155.3 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001
C:Accession: H88451
R:anonymouse, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
A:Reference number: A75000; MIMD:99069613; PMID:9851916
A:Note: See www.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_elegans/
A:Note: published genome appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999
A:Accession: H88451
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1210 <STO>
A:Cross-references: GB:chr.III; PIDN:AA852704.1; PID:g488200; GSPDB:GN00021; CESP:ZC155.
A:Note: weak similarity to Bovine synaptotagmin I
C:Genetics:
A:Gene: ZC155.3
A:Map position: 3

```

Query Match	13.6%	Score 85	DB 2	Length 1210
Best Local Similarity	29.2%	Pred. No. 4.3		
Matches	21	Conservative	16	Mismatches 29
				Indels 6
				Gaps 2

```
OY      *EY-KGEKOLPYLDKIKFLVPDHVNSELVIRRLQDNPLOAQFELLVNCNMSWVSIP   94
          ||| |||| | | | | | | | | | | | | | | | | | | | | | |
Db_     347 ERYCKGNKKKHHCISFTKIMPDNDTDDYKLEKASVNLN---FKSNSHIHIGPLSA 401
```

QY	95	IADIYEQEKDED	106
		:::	
Db	402	IAELVDNAYDAD	413

Search completed: August 11, 2003, 10:16:13  
Job time : 40 secs

**THIS PAGE BLANK (U&PTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 09:50:23 ; Search time 24 Seconds

(without alignments)  
237.093 Million cell updates/sec

Title: US-09-904-603-1  
Perfect score: 625  
Sequence: 1 MPSDRPFKORSRFAHCKEV.....EKDDGFLYVYVASQETFGF 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	ID	Description
1	555	88.8	125	MP3A_HUMAN
2	516	82.6	119	MP13_BOVIN
3	516	82.6	124	MP3B_HUMAN
4	516	82.6	124	MP13_MOUSE
5	516	82.6	141	MP13_RAT
6	381.5	61.0	130	YHM6_CAEEL
7	248	39.7	117	GEF2_HUMAN
8	222	35.5	184	STRP_LACB1
9	216	34.6	117	APG8_YEAST
10	194	31.0	123	YOD9_CAEEL
11	174.5	11.9	660	DNL1_CHLMU
12	73	11.7	260	DH10_ARATH
13	73	11.7	746	RNG3_SCHRO
14	72.5	11.6	224	YPOG_BACSU
15	72.5	11.6	682	NPH1_CAEEL
16	72	11.5	140	APGB_HUMAN
17	72	11.5	141	APGB_MOUSE
18	72	11.5	483	Y045_MYGE
19	72	11.5	1046	POL_FENVI
20	71	11.4	392	KBL_BACSU
21	71	11.4	396	YG78_METJA
22	71	11.4	558	CSG_METJA
23	71	11.4	2048	RRPL_SENDE
24	70.5	11.3	278	V751_METJA
25	70.5	11.3	1085	RBP2_BOVIN
26	70	11.2	718	SL53_BOVIN
27	70	11.2	1272	FMK2_MOUSE
28	70	11.2	2228	RRPL_MOUSE
29	70	11.2	2228	RRPL_SENDE
30	70	11.2	2228	RRPL_SENDE
31	69.5	11.1	592	SYD_THETN
32	69	11.0	511	MATK_WHEAT
33	69	11.0	704	T2D4_DROME

34	69	11.0	1311	FMK2_HUMAN	P51816	homo sapien
35	68.5	11.0	453	SYS_ARCPO	O28244	archaeoglobus
36	68	10.9	113	RS6_HICAP	O8K918	bacteria ap
37	68	10.9	587	CN39_HUMAN	O80117	homo sapien
38	68	10.9	1616	P201_MYGE	Q49429	mycoplasma
39	68	10.9	1741	RPO1_GIALA	P25202	gialdia lam
40	67.5	10.8	257	SPHR_SYNP7	P39663	synochrocyce
41	67.5	10.8	288	Y489_RICPR	Q92457	ticketerisia
42	67.5	10.8	1189	POL_NAEVM	P10272	batodon endo
43	67	10.7	511	MATK_HORDV	Q90233	hordus thot
44	67	10.7	511	MATK_HORDV	P17156	hordus thot
45	67	10.7	718	SL53_CANPA	P31637	canis famli

## ALIGNMENTS

RESULT 1  
MP3A\_HUMAN STANDARD: PRI: 125 AA.  
AC 28-FEB-2003 (rel. 41, Created)  
DT 28-FEB-2003 (rel. 41, Last sequence update)  
DT 28-FEB-2003 (rel. 41, Last annotation update)  
DE Microtubule-associated proteins 1A/1B light chain 3A (MAP1A/MAP1B L1)  
DE A) (MAP1A/1B light chain 3 A).  
GN MAP1A3A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN 11)  
RP SEQUENCE FROM N.A.  
RX MEDLINE-21638749; PubMed-11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.P., Jones M., Stavarides G., Almeida J.P., Babbage A.K., Baggott P., Bailey J., Barlow K.F., Bates K.N., Beare D.M., Beare D.M., Beasley G.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Cole C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.K., Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Gartner J., Graham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall P.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson G.M., Johnson B., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Lehaeslahti M.H., Levesha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McMahon L.J., McLeay K., Murray A.A., Oliver K., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Philmore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Selva H.K., Shownkeen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Stilson J.D., Swann R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Troman A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beek S., Rogers J.  
RT \*The DNA sequence and comparative analysis of human chromosome 20 \*  
RL Nature 414:865-871(2001).  
CC -!- FUNCTION: STRUCTURAL PROTEIN INVOLVED IN THE FILAMENTOUS CROSS-BRIDGING BETWEEN MICROTUBULES AND OTHER SKELETAL ELEMENTS (by similarity).  
CC -!- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, L1, L2 AND L3, CAN ASSOCIATE WITH MAP1A AND MAP1B PROTEINS (by similarity).  
CC -!- SIMILARITY: BELONGS TO THE MAP1 L3 / GEF2 FAMILY.  
CC This SWISS-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AL18520; CAC14078.1; -  
DR Genew: HGNC:6838; MAP1LC3A.  
DR MIM: 601242; -  
DR InterPro: IPR004241; MAP1\_LC3.  
DR Pfam: PF02991; MAP1\_LC3; 1.  
SQ SEQUENCE 125 AA; 14493 MW; 097697B5424FC425 CRC64;

Query Match 88.8%; Score 555; DB 1; Length 125;  
Best Local Similarity 100.0%; Pred. No. 4.1e-49;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 ADRCKEVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPHNMSELYKIRRLQ 73  
DB 18 ADRCKEVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPHNMSELYKIRRLQ 77  
QY 74 NPTQAFLLVNOHSMVSVSTPIADIEQEKDEGFLYVYASQETFG 121  
DB 78 NPTQAFLLVNOHSMVSVSTPIADIEQEKDEGFLYVYASQETFG 125

## RESULT 2

MP3\_BOVIN STANDARD; PRT; 119 AA.  
ID MP3\_BOVIN  
AC 041515;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Microtubule-associated proteins 1A/1B light chain 3 (MAP1A/MAP1B LC3)  
DE (MAP1A/1B light chain 3) (fragment).  
GN MAP1LC3 OR MAP1ALC3.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP TISSUE=Brain;  
RC MEDLINE=94209331; PubMed=7908909;  
RA Mann S.S., Hammarback J.A.;  
RT "Molecular characterization of light chain 3. A microtubule binding  
subunit of MAP1A and MAP1B."  
RL J. Biol. Chem. 269:11492-11497(1994).  
RN [2]  
RP SEQUENCE OF 5-119 FROM N.A.  
RX MEDLINE=98216781; PubMed=9557703;  
RA Meyers G., Stoll D., Gunn M.;  
RT "Insertion of a sequence encoding light chain 3 of microtubule-  
associated proteins 1A and 1B in a pestivirus genome: connection with  
virus cytopathogenicity and induction of lethal disease in cattle."  
RL J. Virol. 72:4139-4148(1998).  
CC -1- FUNCTION: STRUCTURAL PROTEIN INVOLVED IN THE FILAMENTOUS  
CROSS-BRIDGING BETWEEN MICROTUBULES AND OTHER SKELETAL ELEMENTS.  
CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE  
WITH MAP1A AND MAP1B PROTEINS.  
CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.  
CC -1- CAUTION: REF.2 SEQUENCE ORIGINATES FROM STRAIN JACP OF PESTIVIRUS  
TYPE 1 WHICH SEEMS TO CONTAIN A CELLULAR INSERTION OF PART OF  
THE BOVINE HOST MAP1ALC3 GENE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U80885; AAB72082.1; ALT\_SEQ.  
DR InterPro: IPR004241; MAP1\_LC3.

DR Pfam: PF02991; MAP1\_LC3; 1.  
KW Microtubules.

FT CONFLICT 5 5 T -> P (IN REF. 1).  
FT CONFLICT 11 11 T -> S (IN REF. 1).  
FT CONFLICT 13 15 EDR -> DDV (IN REF. 1).  
FT CONFLICT 17 20 EDVR -> KEVO (IN REF. 1).  
FT NON\_TER 119 119  
SQ SEQUENCE 119 AA; 13998 MW; 72A21397AB2626759 CRC64;

Query Match 82.6%; Score 516; DB 1; Length 119;  
Best Local Similarity 83.2%; Pred. No. 3.3e-45;  
Matches 99; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 2 PSDFRFXRSPADRCKEVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPHNM 61  
DB 1 PSDFRFXRSPADRCKEVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPHNM 60  
QY 62 ELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIEQEKDEGFLYVYASQETFG 120  
DB 61 ELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIEQEKDEGFLYVYASQETFG 119

## RESULT 3

MP3B\_HUMAN STANDARD; PRT; 124 AA.  
ID MP3B\_HUMAN  
AC 096208;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Microtubule-associated proteins 1A/1B light chain 3B (MAP1A/MAP1B LC3  
B) (MAP1A/1B light chain 3 B).  
GN MAP1LC3B OR MAP1ALC3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP TISSUE=Lung carcinoma;  
RC TISSUE=Phaeochromocytoma;  
RA Pietas A., Petersen I., Schlens K., Petersen S.;  
RT "Human cDNA of MAP1A/1B LC3 gene which is downregulated in lung  
carcinoma cells."  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Phaeochromocytoma;  
RA Peng Y., Li Y., Gu W., Xu S., Han Z., Fu G., Chen Z.;  
RT "A novel gene expressed in human pheochromocytoma."  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Kikuchi R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,  
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;  
RT "NEDO human cDNA sequencing project."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: STRUCTURAL PROTEIN INVOLVED IN THE FILAMENTOUS CROSS-  
BRIDGING BETWEEN MICROTUBULES AND OTHER SKELETAL ELEMENTS (BY  
SIMILARITY).  
CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE  
WITH MAP1A AND MAP1B PROTEINS (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF303888; AAG23182.1; -  
DR EMBL: AF183417; AAG09686.1; -

```

DR EMBL: AK025556; BAB15169.1;
DR Genew: H6NC:1352; MAP1LC3B.
DR InterPro: IPR004241; MAP1LC3.
DR Pfam: PF02991; MAP1LC3; 1.
KW Microtubules.
FT INITMET 0 0 BY SIMILARITY.
SQ SEQUENCE 124 AA: 14557 MW: BB147B1C5DA3A53F CRC64:
Query Match 82.6%; Score 516; DB 1; Length 124;
Best Local Similarity 82.4%; Pred. No. 3,4e-45;
Matches 98; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 2 PSDFPFKQKRSFADCKEYVQIRDPHSKIPVILIERKKGKQLPVLDTKFLVPHVYNS 61
   1 PSEKTFKQKRSFEGKVEIVHLIRHQPTKIPVILIERKKGKQLPVLDTKFLVPHVYNS 60
DB 1 ELIKTIKRRQLNANQAFLLVNGHSMVSVSTPISEYSESEKDEDFGLVWYVASQETFG 119
DB 61 ELIKTIKRRQLNANQAFLLVNGHSMVSVSTPISEYSESEKDEDFGLVWYVASQETFG 119

RESULT 4
MAP1_MOUSE
ID MAP1_MOUSE STANDARD: PRT: 124 AA.
AC Q6GQV6; Q6DIR0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Microtubule-associated proteins 1A/1B light chain 3 (MAP1A/MAP1B LC3)
DE (MAP1A/1B light chain 3).
DE MAP1LC3 OR MAP1ALC3.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo, Heart, and Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamaka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Furuno M., Aono H., Baldarelli P., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RA Functional annotation of a full-length mouse cDNA collection.";
RI Nature 409:685-690(2001).
CC -1- FUNCTION: STRUCTURAL PROTEIN INVOLVED IN THE FILAMENTOUS CROSS-
CC BRIDGING BETWEEN MICROTUBULES AND OTHER SKELETAL ELEMENTS (BY
CC SIMILARITY).
CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAP1A AND MAP1B PROTEINS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcements/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF255953; AAL83723.1;
DR EMBL: AK002795; BAB22364.1;
DR EMBL: AK003106; BAB22569.1;
DR EMBL: AK003205; BAB22641.1;
DR EMBL: AK003588; BAB22855.1;
DR EMBL: AK012604; BAB28350.1;
DR MGD: MGI:1914694; Map1lc3.
DR InterPro: IPR004241; MAP1LC3.
DR Pfam: PF02991; MAP1LC3; 1.
KW Microtubules.
FT INITMET 0 0 BY SIMILARITY.
SQ SEQUENCE 88 AA: 14486 MW: 520E4FF2AA93B78D CRC64:
Query Match 82.6%; Score 516; DB 1; Length 124;
Best Local Similarity 82.4%; Pred. No. 3,4e-45;
Matches 98; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 2 PSDFPFKQKRSFADCKEYVQIRDPHSKIPVILIERKKGKQLPVLDTKFLVPHVYNS 61
   1 PSEKTFKQKRSFEGKVEIVHLIRHQPTKIPVILIERKKGKQLPVLDTKFLVPHVYNS 60
DB 1 ELIKTIKRRQLNANQAFLLVNGHSMVSVSTPISEYSESEKDEDFGLVWYVASQETFG 119
DB 61 ELIKTIKRRQLNANQAFLLVNGHSMVSVSTPISEYSESEKDEDFGLVWYVASQETFG 119

RESULT 5
MAP1_RAT
ID MAP1_RAT STANDARD: PRT: 141 AA.
AC Q6Z6Z5;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Microtubule-associated proteins 1A/1B light chain 3 (MAP1A/MAP1B LC3)
DE (MAP1A/1B light chain 3).
DE MAP1LC3 OR MAP1ALC3.
GN Rattus norvegicus (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RA [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94209331; PubMed 7908909;
RA Mann S.S., Hammarback J.A.;
RT "Molecular characterization of light chain 3. A microtubule binding
RT subunit of MAP1A and MAP1B.";
RI J. Biol. Chem. 269:11492-11497(1994).
CC -1- FUNCTION: STRUCTURAL PROTEIN INVOLVED IN THE FILAMENTOUS
CC CROSS-BRIDGING BETWEEN MICROTUBULES AND OTHER SKELETAL ELEMENTS.
CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAP1A AND MAP1B PROTEINS.
CC -1- TISSUE SPECIFICITY: ABUNDANT ONLY IN NEURONS. DETECTED IN TESTES.
CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcements/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U05784; AAA20645.1;
DR PIR: A53624; A53624.
DR InterPro: IPR004241; MAP1LC3.
DR Pfam: PF02991; MAP1LC3; 1.

```

KW Microtubules. 0 BY SIMILARITY.  
 FT INIT\_MBT 141 AA: 16263 MW: F14E9D788D03D65D CRC64:  
 SQ SEQUENCE 141 AA: 16263 MW: F14E9D788D03D65D CRC64:  
 Query Match 82.68: Score 516: DB 1: Length 141:  
 Best Local Similarity 82.48: Pred. No. 4e-45:  
 Matches 98: Conservative 12: Mismatches 9: Indels 0: Gaps 0:

QY 2 PSDRPFQRRSFARCKEVOQIRDPHSKIPVIERKYGKQLVLDKTKFLVDPDHVMS 61  
 1 PSEKTFQRRSFARCKEVOQIRDPHSKIPVIERKYGKQLVLDKTKFLVDPDHVMS 60  
 Db 61 ELIKIRRLQNLNANOAFLLVNGHSMVSTPISEYVESRDEDEGLVYVYASQETFG 119

QY 62 ELVKIIRRLQNLNANOAFLLVNGHSMVSTPIADIYDEKDEDEGLVYVYASQETFG 120  
 1 ELVKIIRRLQNLNANOAFLLVNGHSMVSTPIADIYDEKDEDEGLVYVYASQETFG 119  
 Db 61 ELIKIRRLQNLNANOAFLLVNGHSMVSTPISEYVESRDEDEGLVYVYASQETFG 119

RESULT 6  
 YHM6\_CAEEL STANDARD; PRT; 130 AA.  
 ID YHM6\_CAEEL  
 AC Q23536:  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 15.1 kDa protein ZK593.6 in chromosome IV.  
 GN ZK593.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditidae;  
 CC Rhabditidae; Peleoderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RA McMuray A.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z69385; CAAG3421.1;  
 DR PIR: T27920; T27920.  
 DR WormPep: ZK593.6; CE06627.  
 DR InterPro: IPR004241; MAP1\_LC3.  
 DR Pfam: PF02991; MAP1\_LC3; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 130 AA: 15095 MW: 347518DF8C8D3154 CRC64:  
 Query Match 61.08: Score 381.5: DB 1: Length 130:  
 Best Local Similarity 58.38: Pred. No. 1.3e-31:  
 Matches 70: Conservative 27: Mismatches 20: Indels 3: Gaps 1:

QY 1 MPSPDRPFQRRSFARCKEVOQIRDPHSKIPVIERKYGKQLVLDKTKFLVDPDHVMS 60  
 1 PSEKTFQRRSFARCKEVOQIRDPHSKIPVIERKYGKQLVLDKTKFLVDPDHVMS 60  
 Db 14 VPS---FKERRPFERKQKDYBEIRSQOPKNVPIIEERFDEGRSLPLMDRCKFLVPEHITY 70  
 QY 61 SELVKIIRRLQNLNANOAFLLVNGHSMVSTPIADIYDEKDEDEGLVYVYASQETFG 120  
 1 SELVKIIRRLQNLNANOAFLLVNGHSMVSTPIADIYDEKDEDEGLVYVYASQETFG 120  
 Db 71 AELMSIVRRRLQNLNANOAFLLVNGHSMVSTPISEYVESRDEDEGLVYVYASQETFG 119

RESULT 7  
 GEF2\_HUMAN STANDARD; PRT; 117 AA.  
 ID GEF2\_HUMAN  
 AC O08765; Q9DCP8; Q9DQF7;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Ganglioside expression factor 2 (GEF-2) (General protein transport  
 DE factor p16) (GATE-16) (GABA(A) receptor-associated protein-like 2)  
 DE (MAP1 light chain 3 related protein).  
 GN GABARAPL2 OR GEF2 OR PLC3A.  
 OS Homo sapiens (Human).  
 OS Mus musculus (Mouse).  
 OS Rattus norvegicus (Rat), and  
 OS Bos taurus (Bovine).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606, 10090, 10116, 9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Human;  
 RA Storch S., Braulke T.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Human;  
 RA Song H., Peng Y., Yu Y., Fu G., Mao M., Zhang Q., Zhu H., Li G.,  
 RA Luo M., Chen J., Hu R.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A., AND INTERACTION WITH UTK1.  
 RC SPECIES-Human; TISSUE-Frontal cortex;  
 RX MEDLINE=21066693; PubMed=11146101;  
 RA Okazaki N., Yan J., Yuasa S., Ueno T., Komitani E., Masuho Y.,  
 RA Koga H., Muramatsu M.A.;  
 RL "Interaction of the unc-51-like kinase and microtubule-associated  
 RL protein light chain 3 related proteins in the brain: possible role of  
 RL vesicular transport in axonal elongation."  
 RL Brain Res. Mol. Brain Res. 85:1-12(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Human; TISSUE-Brain, and Testis;  
 RA Strausberg R.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Human, and Mouse;  
 RX MEDLINE=21309076; PubMed=11414770;  
 RA Xin Y., Yu L., Chen Z., Zheng L., Fu Q., Jiang J., Zhang P., Gong R.,  
 RA Zhao S.;  
 RL "Cloning, expression patterns, and chromosome localization of three  
 RL human and two mouse homologues of GABA(A) receptor-associated  
 RL protein."  
 RL Genomics 74:408-413(2001).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Rat; TISSUE-Brain;  
 RA Ogura K.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Bovine; TISSUE-Brain;  
 RX MEDLINE=20211392; PubMed=10747018;  
 RA Savit Y., Legesse-Miller A., Porat A., Elazar Z.;  
 RL "GATE-16, a membrane transport modulator, interacts with NSF and the  
 RL Golgi v-SNARE GOS-28."  
 RL EMBO J. 19:1494-1504(2000).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Bovine; TISSUE-Kidney;  
 RX MEDLINE=22326861; PubMed=12438634;  
 RA Becher P., Thiel H.-J., Collins M., Brownlie J., Orlich M.;  
 RL "Cellular sequences in pestivirus genomes encoding gamma-aminobutyric  
 RL Acid (A) receptor-associated protein and Golgi-associated ATPase  
 RL enhancer of 16 kilodaltons."  
 RL J. Virol. 76:13069-13076(2002).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Mouse; STRAIN-C57BL/6J; TISSUE-Hippocampus, and Kidney;







AC Q9PKP2: 16-OCT-2001 (Rel. 40, Created)  
 DI 16-OCT-2001 (Rel. 40, Last sequence update)  
 DI 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA ligase (EC 6.5.1.2) (Polydeoxyribonucleotide synthase [NAD+]).  
 GN LIGA OR TC0423.  
 OS Chlamydia muridarum.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_taxid=83560;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mopn / N199;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,  
 RA White O., Hickey E.R., Peterson J., Ulteback T., Berry K., Bass S.,  
 RA Linher M., Weidman J., Khouri H., Craven R., Bowman C., Dodson R.,  
 RA Gunn M., Nelson W., Deboy K., Kolonay J., McClarty G., Salzberg S.L.,  
 RA Eisen J., Fraser G.M.;  
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia  
 RT pneumoniae AR39.";  
 RI Nucleic Acids Res. 28:1397-1406(2000).  
 CC -1- FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER  
 CC LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-  
 CC STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR  
 CC THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF  
 CC DAMAGED DNA (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: NAD(+) + [deoxyribonucleotide(N) +  
 CC (deoxyribonucleotide(N)+M) = AMP + nicotinamide nucleotide +  
 CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.  
 CC -1- SIMILARITY: Contains 1 BRCT domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AE002309; AAF39279.1; -;  
 DR PIR: A81704; A81704.  
 DR HSSP: 087703; 1B04.  
 DR TIGR: TC0423; -;  
 DR InterPro: IPR001357; BRCT.  
 DR InterPro: IPR004150; DNA\_ligase\_OB.  
 DR InterPro: IPR001679; DNALigase.  
 DR InterPro: IPR000445; HhH.  
 DR InterPro: IPR003583; HhH\_1.  
 DR InterPro: IPR004149; Znf\_DNAligase\_C4.  
 DR Pfam: PF00533; BRCT; 1.  
 DR Pfam: PF01653; DNA\_ligase\_N; 1.  
 DR Pfam: PF03120; DNA\_ligase\_OB; 1.  
 DR Pfam: PF03119; DNA\_ligase\_CBD; 1.  
 DR Pfam: PF00633; HhH; 1.  
 DR Pfam: PF003944; DNAligase; 1.  
 DR SMART: SM00292; BRCT; 1.  
 DR SMART: SM00278; HhH; 1.  
 DR SMART: SM00532; LIGANC; 1.  
 DR TIGRFAMs: TIGR00575; dnlj; 1.  
 DR PROSITE: PS50172; BRCT; 1.  
 DR PROSITE: PS01055; DNA\_LIGASE\_N1; FALSE.NEG.  
 DR PROSITE: PS01056; DNA\_LIGASE\_N2; 1.  
 KW LIGASE; DNA repair; DNA replication; NAD; Complete proteome.  
 FT BINDING: 581 660 BRCT.  
 FT DOMAIN: 114 114 AMP (BY SIMILARITY).  
 FT SEQUENCE: 660 AA; 73784 MW; 06BB89145E723EAD CRC64;  
 Query Match 11.9%; Score 74.5; DB 1; Length 660;  
 Best Local Similarity 24.5%; Pred. No. 7.5;  
 Matches 26; Conservative 16; Mismatches 27; Indels 37; Gaps 3;  
 QY 4 DRPKRSPADKCKEVOIDQDHS-KPIVIERYKGEKQPLVDKTKFLVPHVMNSE 62

DB 27 DRPTISDYSYDMKMEHLEQHEKEMKVPSPITYLQDR-----  
 QY 63 LKTIIRRLQNLPTQAFPLVNGHSMVSVPIADITQEKDEQCF 108  
 DB 67 -----PSGPIVVPESHFMVS-----IANVSLSELEEFF 96  
 RESULT 12  
 DH10\_ARATH STANDARD: 260 AA.  
 ID DH10\_ARATH PR1: 260 AA.  
 AC P42759;  
 DI 01-NOV-1995 (Rel. 32, Created)  
 DI 01-NOV-1995 (Rel. 32, Last sequence update)  
 DI 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Dehydrin ERD10 (low-temperature-induced protein 11145).  
 GN ERD10 OR L1145 OR L1129 OR A11G20450 OR F5M15.21.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosales;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=94348723; PubMed=8069491;  
 RA Kiyosue T., Yamaguchi-Shinozaki K., Shinozaki K.;  
 RT "Characterization of two cDNAs (ERD10 and ERD13) corresponding to  
 RT genes that respond rapidly to dehydration stress in Arabidopsis  
 RT thaliana.";  
 RI Plant Cell Physiol. 35:225-231(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Landsberg erecta; TISSUE:leaf;  
 RX MEDLINE=96046757; PubMed=7579189;  
 RA Wellin B.V., Olson A., Palva E.T.;  
 RT "Structure and organization of two closely related low-temperature-  
 RT induced dhn/lea/rab-like genes in Arabidopsis thaliana L. Heynh.";  
 RI Plant Mol. Biol. 29:391-395(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=1130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Decker K.,  
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam H.,  
 RA Langan-Hooper S., Lee A., Lee J.M., Leuz G.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti K., Marziani A.,  
 RA Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.L.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley I.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Tomin M.J., Town C.D.,  
 RA Ulteback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser G.M., Venter J.C., Davis K.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana.";  
 RI Nature 408:816-820(2000).  
 RL Nature 408:816-820(2000).  
 RN [4]  
 RP SEQUENCE OF 56-260 FROM N.A.  
 RC STRAIN=cv. Landsberg erecta; TISSUE:leaf;  
 RX MEDLINE=95035985; PubMed=7948863;  
 RA Wellin B.V., Olson A., Nylander M., Palva E.T.;  
 RT "Characterization and differential expression of dhn/lea/rab-like  
 RT genes during cold acclimation and drought stress in Arabidopsis  
 RT thaliana.";  
 RI Plant Mol. Biol. 26:131-144(1994).  
 CC -1- TISSUE SPECIFICITY: IN STEMS, CAULINE LEAVES, ROOTS AND FLWERS,  
 CC LOW LEVELS FOUND IN MATURING SEEDS. ABSENT IN DRY SEEDS.

CC -1- INDUCTION: BY DEHYDRATION, COLD STRESS AND ABSISIC ACID (ABA).  
 CC INDUCTION AFTER 10 H. COLD STRESS INDUCTION PEAKS AT 1 H AND 5 H  
 CC MAXIMUM AFTER 10 H. COLD STRESS INDUCTION PEAKS AT 1 H AND 5 H  
 CC AFTER START OF COLD EXPOSURE.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEHYDRIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D17714: BAA04568.1; -  
 DR EMBL: X90958: CAA62448.1; -  
 DR EMBL: AC027665: AAF79613.1; -  
 DR EMBL: X77614: CAA54705.1; -  
 DR PIR: S60480: S60480.  
 DR InterPro: IPR000167: Dehydrin.  
 DR Pfam: PF00257: dehydrin.1.  
 DR PROSITE: PS00315: DEHYDRIN\_1. 1.  
 DR PROSITE: PS00823: DEHYDRIN\_2. 2.  
 KW Dehydrin; Repeat; Multigene family.  
 FT DOMAIN 71 74 POLY-GLU.  
 FT DOMAIN 88 92 POLY-GLU.  
 FT DOMAIN 104 112 POLY-SER.  
 FT DOMAIN 121 129 POLY-LYS.  
 FT DOMAIN 183 186 POLY-GLU.  
 FT REPEAT 184 247 2 X 21 AA REPEATS, LYS-RICH.  
 FT REPEAT 227 247 2.  
 FT REPEAT 256 256 H -> G (IN REF. 3).  
 FT CONFLICT 144 144 Q -> R (IN REF. 3).  
 SQ SEQUENCE 260 AA: 29548 MW: 31866CB6B4A1A5A CRC64;

Query Match 11.7%; Score 73; DB 1; Length 260;  
 Best Local Similarity 24.8%; Pred. No. 3.5;  
 Matches 28; Conservative 21; Mismatches 34; Indels 30; Gaps 5;

QY 8 KQRRSPADRCKEVQQRDQ-HSKIKYV-----IERKGEKQ--LPVLDTKFLVVD 56  
 DB 141 EENQGVMDVIRKKEFPKPGDDVPPVTMPAPHSVDEHKKEEKEKGFMRKIKELPG 200

QY 57 HWMSSELVRIIRRLQNLPTQAFELLVNQHWSVSTPIADIYDEKEDDGL 109  
 DB 201 HSKKPEDSOV-----VNTTPIVETATPIADIPEEK---GFM 234

RESULT 13  
 RN3\_SCHPO STANDARD; PRT: 746 AA.  
 AC 074994:  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ring assembly protein 3.  
 GN RN3 OR SPC613.04C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.  
 RC STRAIN=972;  
 RX MEDLINE=20312839; PubMed=10852821;  
 RA Wong K.C.Y., Nagy N.I., Iino Y., Yamamoto M., Balasubramanian M.K.;  
 RT "Fission yeast Rng3p: an UCS-domain protein that mediates myosin II  
 RL assembly during cytokinesis.";  
 RN J. Cell Sci. 113:2421-2432(2000).  
 RP SEQUENCE FROM N.A.

RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., Mcdonald S., Mclean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford J., Rutter S., Saunders R., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Welter J., Wolckart G., Aert R., Robben J., Grymonprez B.,  
 RA Welfens I., Vanstreets E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Srpakowski G.V., Ussey D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -1- FUNCTION: Essential for actinomyosin ring assembly during  
 CC cytokinesis. Has a role in conjunction with F-actin, in  
 CC assembling myosin II-containing proteins, such as myo2, at the  
 CC division site.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; associated with the  
 CC actinomyosin ring.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ011773: CAA09767.1; -  
 DR EMBL: AL031644: CAA21056.1; -  
 DR PIR: T43648: T43648.  
 DR GeneDB\_Spomb: SPC613.04C; -  
 KW Cell division; Septation.  
 SQ SEQUENCE 746 AA: 84471 MW: A0158CEB110D71E CRC64;

Query Match 11.7%; Score 73; DB 1; Length 746;  
 Best Local Similarity 23.9%; Pred. No. 12;  
 Matches 28; Conservative 25; Mismatches 44; Indels 20; Gaps 6;

QY 7 FKQRSFDR-CKEVOQRDQPSKIPYIE-----RKGEKQLPVLDTKFLVVDHNM 60  
 DB 291 FKRCRAIIRKKESEF--IKDNTSRVAILISKKAFOYLSQVPIPTL---KLLOSKEYD 345

QY 61 SELVRIIRRLQNLPTQAF-----LLVNQHWSVSTPIADIYDEKEDDGLVWY 113  
 DB 346 SVLEALNQSSTLGFVKQLIADDSCLNLSKLLDTNISP-----DASSIATPIY 397

RESULT 14  
 YPJG\_BACSU STANDARD; PRT: 224 AA.  
 AC P42981:  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein ypjG.  
 GN YPJG OR JOJG.  
 OS Bacillus subtilis.

```

603 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
604 NCBI_TaxID=1423;
605 [1]
606 SEQUENCE FROM N.A.
607 STRAIN=168 / Maltodug;
608 MEDLINE=96349105; PubMed=8760912;
609 Sotokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
610 Serrif P.;
611 "Sequence analysis of the Bacillus subtilis chromosome region between
612 the srcA and kds loci cloned in a yeast artificial chromosome.";
613 Microbiology 142:2005-2016(1996).
614 [2]
615 SEQUENCE FROM N.A.
616 STRAIN=168;
617 MEDLINE=98044033; PubMed=9384377;
618 Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
619 Azevedo V., Berrtero M.G., Bessieres P., Borotini A., Borcieri S.,
620 Borris R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,
621 Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
622 Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
623 Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
624 Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
625 Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
626 Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
627 Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
628 Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
629 Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
630 Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
631 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
632 Lee S.M., Levine A., Liu H., Masuda S., Malet C., Medigue C.,
633 Medina N., Melillo R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
634 Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
635 Parren V., Pohl T.M., Portelle D., Portollik S., Prescott A.M.,
636 Prensier E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
637 Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
638 Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,
639 Sekiguchi J., Sekowska A., Serrif S.J., Serrif P., Shin B.S., Soldo B.,
640 Sotokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
641 Takeuchi M., Tanakoshi A., Tanaka T., Tappara P., Tognoni A.,
642 Tosato V., Uchiyama S., Vandebol M., Vannier F., Vassartoli A.,
643 Viari A., Wambuit R., Wedler E., Wedler H., Weitzenger T.,
644 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
645 Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
646 "The complete genome sequence of the gram-positive bacterium Bacillus
647 subtilis.";
648 Nature 390:249-256(1997).
649 [3] SIMILARITY: TO F.COLI YAIS.
650 -----
651 This SWISS-PROT entry is copyright. It is produced through a collaboration
652 between the Swiss Institute of Bioinformatics and the EMBL outstation -
653 the European Bioinformatics Institute. There are no restrictions on its
654 use by non-profit institutions as long as its content is in no way
655 modified and this statement is not removed. Usage by and for commercial
656 entities requires a license agreement (See http://www.isb-sib.ch/announce/
657 or send an email to license@isb-sib.ch).
658 -----
659 EMBL: 138424; AAA92876.1; -
660 EMBL: 147709; AAB38444.1; -
661 EMBL: 299115; CAB14163.1; -
662 EMBL: F69937; F69937.
663 Subtilisin, Bg11212; ypfG.
664 InterPro: IPR003737; DUF158.
665 Pfam: PF02585; DUF158. 1.
666 Hypothetical protein; Complete proteome.
667 SEQUENCE 224 AA; 24806 MW; 95984EAF3F47090 CMC64;
668
669 Query Match 11.6%; Score 72.5; DB 1; Length 224;
670 Best Local Similarity 24.8%; Pred. No. 3.3;
671 Matches 29; Conservative 22; Mismatches 31; Indels 35; Gaps 6;
672
673 QY 24 KDHPSK-----IVTTERYKGEKQLPVLDKTK-----FLVPDHV-NMSEL 63
674 Db 107 KDRHPDHGNAALVEAIFSAIGIKRDEKSLPAHKVSKYVYWMINGFHQDFVIDISDT 166

```

```

675 QY 64 VKIIRRLQJNPTQAFLLVNHGSMVSSTPIADYEGEKDEKDEFLVNVASQETH; 129
676 Db 167 IE-AKKOSLNATYSQETH-----PSKDSVSTPLI-----NGTIEVAKRKVIY 204
677
678 RESULT 15
679 NPH1_CAEEL
680 ID NPH1_CAEEL STANJARD; PRI: 682 AA.
681 AC 017972; Q21570;
682 DT 28-FEB-2003 (rel. 41, Created)
683 DT 28-FEB-2003 (rel. 41, Last sequence update)
684 DE 28-FEB-2003 (rel. 41, Last annotation update)
685 DE Nephrocytin-1 like protein.
686 GN M28.7
687 OS Caenorhabditis elegans.
688 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
689 Rhabditiidae; Pelodierinae; Caenorhabditis.
690 OX NCBI_TaxID=6239;
691 [1]
692 SEQUENCE FROM N.A.
693 STRAIN=Bristol N2;
694 Gardner A.;
695 Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
696 [3] SIMILARITY: Contains 1 SH3 domain.
697 [3] SIMILARITY: BELONGS TO THE NEPHROCYTIN 1 FAMILY.
698 -----
699 This SWISS-PROT entry is copyright. It is produced through a collaboration
700 between the Swiss Institute of Bioinformatics and the EMBL outstation -
701 the European Bioinformatics Institute. There are no restrictions on its
702 use by non-profit institutions as long as its content is in no way
703 modified and this statement is not removed. Usage by and for commercial
704 entities requires a license agreement (See http://www.isb-sib.ch/announce/
705 or send an email to license@isb-sib.ch).
706 -----
707 EMBL: 249911; GAA90133.1; -
708 PIR: T23813; T23813.
709 Wormpep: M28.7; CE16278.
710 InterPro: IPR001452; SH3.
711 Pfam: PF00018; SH3; 1.
712 SMART: SM00326; SH3; 1.
713 DR PROSITE: PS50002; SH3; 1.
714 KW Hypothetical protein; Coiled coil; SH3 domain.
715 FT DOMAIN 10 37 COILED COIL (POTENTIAL).
716 FT DOMAIN 43 100 COILED COIL (POTENTIAL).
717 FT DOMAIN 180 200 COILED COIL (POTENTIAL).
718 FT DOMAIN 215 275 SH3.
719 FT DOMAIN 155 173 GLN/PRO-RICH.
720 SEQUENCE 682 AA; 78795 MW; F30CCF8B4B552D CMC64;
721
722 Query Match 11.6%; Score 72.5; DB 1; Length 682;
723 Best Local Similarity 24.3%; Pred. No. 12;
724 Matches 25; Conservative 16; Mismatches 43; Indels 17; Gaps 4;
725
726 QY 10 RRFADFQREYQVQIRQDHPKRI--PVIIKKYKGRQLPVLDKTK-----FLVPDHV-NMSEL 64
727 Db 43 RHFVPCQQLHRLQDLDRHNRKLEKAKIDETVKKKNALQJDKLQRLTALSPKEQSLSPS 102
728
729 QY 64 VKIIRRLQJNPTQAFLLVNHGSMVSSTPIADYEGEKDEKDEFLVNVASQETH; 105
730 Db 103 VSVDSQSEERK-----KMAAIGRKSTMYNDSESD 134

```

Search completed: August 11, 2003, 10:14:45  
Job time : 26 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 10:08:18 ; Search time 95 Seconds  
(without alignments)  
328.678 Million cell updates/sec

Title: US-09-904-603-1

Perfect score: 625  
Sequence: 1 MPSPRFXKGRSFADRCKEY.....EKDEDGLYVYASQRTFGF 121

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL\_23: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_plant: \*  
10: sp\_prodent: \*  
11: sp\_virus: \*  
12: sp\_vertebrate: \*  
13: sp\_unclassified: \*  
14: sp\_virus: \*  
15: sp\_bacteriap: \*  
16: sp\_bacteriap: \*  
17: sp\_archaeap: \*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625	100.0	121	4 Q9BXW5	Q9BXW5 homo sapien
2	625	100.0	121	11 Q91VR7	Q91VR7 mus musculu
3	625	100.0	170	4 Q9H491	Q9H491 homo sapien
4	622	99.5	121	11 Q9DC74	Q9DC74 mus musculu
5	359	57.4	147	4 Q9BXW4	Q9BXW4 homo sapien
6	242	38.7	339	12 Q8B512	Q8B512 bovine vira
7	234	37.4	116	10 Q8H715	Q8H715 phytochthor
8	234	37.4	121	10 Q8VYK7	Q8VYK7 arabidopsi
9	230	36.8	120	10 Q8LBA9	Q8LBA9 arabidopsi
10	227	36.3	120	10 Q9SL04	Q9SL04 arabidopsi
11	227	36.3	122	10 Q945K6	Q945K6 arabidopsi
12	227	36.3	122	10 Q8LEW4	Q8LEW4 arabidopsi
13	226	36.2	122	10 Q8S928	Q8S928 arabidopsi
14	225	36.0	122	10 Q8S926	Q8S926 arabidopsi
15	224	35.8	121	3 Q94272	Q94272 schizosach
16	223	35.7	115	10 Q9LRP7	Q9LRP7 arabidopsi

17	222	35.5	121	10 Q815F9	Q815F9 clect arid
18	221	35.4	121	3 Q8M2Y7	Q8M2Y7 neopostera
19	221	35.4	121	3 Q81282	Q81282 postospira a
20	219	35.0	107	10 Q65447	Q65447 arabidopsi
21	219	35.0	122	10 Q9XEB5	Q9XEB5 arabidopsi
22	217	34.7	119	10 Q8S925	Q8S925 arabidopsi
23	217	34.7	119	10 Q8S927	Q8S927 arabidopsi
24	217	34.7	121	10 Q912Z9	Q912Z9 arabidopsi
25	208	34.3	111	6 Q8H1B6	Q8H1B6 bos taurus
26	208	34.3	129	10 Q9S0U5	Q9S0U5 arabidopsi
27	203	32.5	117	4 Q9H0R8	Q9H0R8 homo sapien
28	203	32.5	117	11 Q9J197	Q9J197 caria porce
29	203	32.5	239	12 Q8B513	Q8B513 bovine vira
30	201.5	32.2	149	10 Q80708	Q80708 arabidopsi
31	200	32.0	125	5 Q9N1J4	Q9N1J4 pichia past
32	199	31.8	117	11 Q8R4B8	Q8R4B8 mus musculu
33	194	31.0	117	4 Q9HYV6	Q9HYV6 homo sapien
34	186	29.8	121	5 Q9W2S2	Q9W2S2 diosyllia
35	184	29.4	124	5 Q81JK2	Q81JK2 plasmodium
36	183	29.3	117	4 Q951G5	Q951G5 homo sapien
37	183	29.3	117	6 Q8KK68	Q8KK68 myctolagus
38	183	29.3	117	11 Q9Q017	Q9Q017 mus musculu
39	181	29.0	117	11 Q912T6	Q912T6 mus musculu
40	177	28.3	120	5 Q9VEG5	Q9VEG5 diosyllia
41	172	27.5	122	13 Q9DFN7	Q9DFN7 allitichys
42	140	22.4	165	10 Q23496	Q23496 arabidopsi
43	92	14.7	829	17 Q80499	Q80499 pyrococcus
44	88.5	14.2	512	8 Q9GHG5	Q9GHG5 colididia )
45	85	13.6	841	17 Q59612	Q59612 pyrococcus

# ALIGNMENTS

## RESULT 1

Q9BXW5 PREDIMINARY: FR: 121 AA.  
ID Q9BXW5: AC Q9BXW5:  
BT 01-JUN-2001 (TREMBLrel, 17, Created)  
DT 01-JUN-2001 (TREMBLrel, 17, Last sequence update)  
DI 01-OCT-2002 (TREMBLrel, 22, Last annotation update)  
DE MAPI light chain 3-like protein 1 (Hypothetical protein).  
GN DKFZP761L0515.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yu L.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=uterus;  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Amygdala;  
RA Ansgore W., Winkler U., Mewes H.W., Weil R., Wiemann S.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL: AF276658; AAK35151.1; -;  
DR EMBL: BC015810; AAH15810.1; -;  
DR EMBL: AL833855; CAD38714.1; -;  
DR InterPro: IPR004241; MAPI\_LC3;  
DR Pfam: PF02991; MAPI\_LC3; 1;  
KW Hypothetical protein.  
SQ SEQUENCE 121 AA: 14272 MW: 4801FRRRF789ZAF3 CR764:

Query Match 100.0%; Score 625; DB 4; Length 121;  
Best Local Similarity 100.0%; Prod. No. 176-56;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPDRPFKORRSFADCKEVOQIRDOHPKIPVIERKYGKQQLPVLDKTKFLVDPDHVM 60  
DB 1 MSPDRPFKORRSFADCKEVOQIRDOHPKIPVIERKYGKQQLPVLDKTKFLVDPDHVM 60  
QY 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYQEKDEGDFLYMYASQETFG 120  
DB 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYQEKDEGDFLYMYASQETFG 120  
QY 121 F 121  
DB 121 F 121

## RESULT 2

Q91VR7 PRELIMINARY; PRT: 121 AA.  
AC Q91VR7;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE MAPI light chain 3-like protein 1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Breast tumor;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC010596; ANH10596.1;  
DR InterPro: IPR004241; MAP1\_LC3;  
DR Pfam: PF02991; MAP1\_LC3; 1.  
SQ SEQUENCE 121 AA: 14272 MW: 48C1FBEBF7892AF3 CRC64;

Query Match Best Local Similarity 100.0%; Score 625; DB 11; Length 121;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPDRPFKORRSFADCKEVOQIRDOHPKIPVIERKYGKQQLPVLDKTKFLVDPDHVM 60  
DB 1 MSPDRPFKORRSFADCKEVOQIRDOHPKIPVIERKYGKQQLPVLDKTKFLVDPDHVM 60  
QY 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYQEKDEGDFLYMYASQETFG 120  
DB 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYQEKDEGDFLYMYASQETFG 120  
QY 121 F 121  
DB 121 F 121

## RESULT 3

Q9H491 PRELIMINARY; PRT: 170 AA.  
AC Q9H491;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE BA346K17.1.2 (Novel protein similar to MAP1LC3 (Microtubule-associated  
protein 1A/1B light chain 3) from Rat, Isoform 2) (Fragment).  
GN BA346K17.1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Saira H.;  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL118520; CAC14079.1;  
DR InterPro: IPR004241; MAP1\_LC3;  
DR Pfam: PF02991; MAP1\_LC3; 1.

FT NON TER 1 1  
SQ SEQUENCE 170 AA: 19189 MW: B8C7521EDC6E4823 CRC64;  
Query Match Best Local Similarity 100.0%; Score 625; DB 4; Length 170;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPDRPFKORRSFADCKEVOQIRDOHPKIPVIERKYGKQQLPVLDKTKFLVDPDHVM 60  
DB 50 MSPDRPFKORRSFADCKEVOQIRDOHPKIPVIERKYGKQQLPVLDKTKFLVDPDHVM 109  
QY 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYQEKDEGDFLYMYASQETFG 120  
DB 110 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYQEKDEGDFLYMYASQETFG 169  
QY 121 F 121  
DB 170 F 170

## RESULT 4

Q9DC74 PRELIMINARY; PRT: 121 AA.  
AC Q9DC74;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DE 1010001H21RIK protein.  
GN 1010001H21RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Heart;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Salto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Salto R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Guslinich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
DR EMBL: AK003122; BAB22582.1;  
DR MGI: MGI:1915661; 1010001H21RIK.  
DR InterPro: IPR004241; MAP1\_LC3.  
DR Pfam: PF02991; MAP1\_LC3; 1.  
SQ SEQUENCE 121 AA: 14286 MW: B8C1FBEDA2DF9A9E CRC64;

Query Match Best Local Similarity 99.5%; Score 622; DB 11; Length 121;  
Matches 120; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPDRPFKORRSFADCKEVOQIRDOHPKIPVIERKYGKQQLPVLDKTKFLVDPDHVM 60  
DB 1 MSPDRPFKORRSFADCKEVOQIRDOHPKIPVIERKYGKQQLPVLDKTKFLVDPDHVM 60  
QY 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYQEKDEGDFLYMYASQETFG 120  
DB 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYQEKDEGDFLYMYASQETFG 120









Db 65 VRRRIKLGAEKAFIFV-KNTLPPTAALMSAIEEHKDEGFLYMYSGENTFG 117

## RESULT 14

ID 08S926 PRELIMINARY; PRT; 122 AA.

DT 01-JUN-2002 (TReMBLrel. 21, Created)

DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)

DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)

DE Autophagy 8e.

GN ATAPG8E.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Hanaka H., Noda T., Shirano Y., Sato S., Kato T., Hayashi H.,

RT "Disruption of AtAPG9 indicates autophagy is involved in leaf

senescence and starvation tolerance in higher plant."

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB073179; BAB88391.1; -

DR InterPro: IPR004241; MAPL\_IC3.

DR Pfam: PF02991; MAPL\_IC3; 1.

SO SEQUENCE 122 AA; 13948 MW; 9C8693382DA3A1B CRC64;

Query Match 36.0%; Score 225; DB 10; Length 122;

Best Local Similarity 39.5%; Pred. No. 1.8e-15;

Matches 47; Conservative 29; Mismatches 31; Indels 12; Gaps 4;

Db 7 FKORRSFADRCVQOIRDOHPSKIPYIIERTKGEK-QLPVLDTKFLVDPHVNMSGLYK 65

66 IIRRLQNLPTQAFLLVNOHSMVSVSTP---IADIYEOEKDEGFLYMYVASOETFG 120

65 VIKRKILSPEKAFIFIDE---NLPTAALMSTIYEELKSEDEGFLYMYSGENTFG 118

Db 65 VIKRKILSPEKAFIFIDE---NLPTAALMSTIYEELKSEDEGFLYMYSGENTFG 118

## RESULT 15

ID 094272 PRELIMINARY; PRT; 121 AA.

DT 01-MAY-1999 (TReMBLrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Putative autophagy protein.

GN SPBP8B7.24C.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI\_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RA Beck A., Reinhardt R., Lyne M., Rajandream M.A., Barrell B.G.;

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL032684; CAA21809.1; -

DR GeneDB; SPBP8B7.24C; -

DR InterPro: IPR004241; MAPL\_IC3.

DR Pfam: PF02991; MAPL\_IC3; 1.

SO SEQUENCE 121 AA; 14155 MW; 964BC3EF5FAE983 CRC64;

Query Match 35.8%; Score 224; DB 3; Length 121;

Best Local Similarity 37.7%; Pred. No. 2.2e-15;

Matches 43; Conservative 31; Mismatches 38; Indels 2; Gaps 2;

Db 7 FKORRSFADRCVQOIRDOHPSKIPYIIERTKGEK-QLPVLDTKFLVDPHVNMSGLYK 66

Db 5 FKDDFSEKRRKTESQRIREKYPDIRIPYICKV-DKSDIALDKKKYLVPSDLVGGFVYV 63

QY 67 IIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEOEKDEGFLYMYVASOETFG 120

Db 64 IIRRKILSPEKAFIFIDE-ILPPTAALMSTIYEELKSEDEGFLYMYSGENTFG 116

Search completed: August 11, 2003, 10:15:28  
Job time : 97 secs